Earth Surface Processes and Landforms
Earth Surf. Process. Landforms **28**, 925–950 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/esp.470

TESTING A MODEL FOR PREDICTING THE TIMING AND LOCATION OF SHALLOW LANDSLIDE INITIATION IN SOIL-MANTLED LANDSCAPES

M. CASADEI,[1]* W. E. DIETRICH[1] AND N. L. MILLER[2]

[1] *Department of Earth and Planetary Science, University of California at Berkeley, Berkeley, CA, USA*
[2] *Regional Climate Center – Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Received 10 January 2002; Revised 6 September 2002; Accepted 3 October 2002

ABSTRACT

The growing availability of digital topographic data and the increased reliability of precipitation forecasts invite modelling efforts to predict the timing and location of shallow landslides in hilly and mountainous areas in order to reduce risk to an ever-expanding human population. Here, we exploit a rare data set to develop and test such a model. In a 1·7 km^2 catchment a near-annual aerial photographic coverage records just three single storm events over a 45 year period that produced multiple landslides. Such data enable us to test model performance by running the entire rainfall time series and determine whether just those three storms are correctly detected. To do this, we link a dynamic and spatially distributed shallow subsurface runoff model (similar to TOPMODEL) to an infinite slope model to predict the spatial distribution of shallow landsliding. The spatial distribution of soil depth, a strong control on local landsliding, is predicted from a process-based model. Because of its common availability, daily rainfall data were used to drive the model. Topographic data were derived from digitized 1 : 24 000 US Geological Survey contour maps. Analysis of the landslides shows that 97 occurred in 1955, 37 in 1982 and five in 1998, although the heaviest rainfall was in 1982. Furthermore, intensity–duration analysis of available daily and hourly rainfall from the closest raingauges does not discriminate those three storms from others that did not generate failures. We explore the question of whether a mechanistic modelling approach is better able to identify landslide-producing storms. Landslide and soil production parameters were fixed from studies elsewhere. Four hydrologic parameters characterizing the saturated hydraulic conductivity of the soil and underlying bedrock and its decline with depth were first calibrated on the 1955 landslide record. Success was characterized as the most number of actual landslides predicted with the least amount of total area predicted to be unstable. Because landslide area was consistently overpredicted, a threshold catchment area of predicted slope instability was used to define whether a rainstorm was a significant landslide producer. Many combinations of the four hydrological parameters performed equally well for the 1955 event, but only one combination successfully identified the 1982 storm as the only landslide-producing storm during the period 1980–86. Application of this parameter combination to the entire 45 year record successfully identified the three events, but also predicted that two other landslide-producing events should have occurred. This performance is significantly better than the empirical intensity–duration threshold approach, but requires considerable calibration effort. Overprediction of instability, both for storms that produced landslides and for non-producing storms, appears to arise from at least four causes: (1) coarse rainfall data time scale and inability to document short rainfall bursts and predict pressure wave response; (2) absence of local rainfall data; (3) legacy effect of previous landslides; and (4) inaccurate topographic and soil property data. Greater resolution of spatial and rainfall data, as well as topographic data, coupled with systematic documentation of landslides to create time series to test models, should lead to significant improvements in shallow landslides forecasting. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: shallow landslides; slope stability; hydrology; digital terrain modelling

INTRODUCTION

Owing to advances in the temporal and spatial resolution in precipitation forecasting (Miller and Kim, 1997a; Seo *et al*., 1999), we are approaching a time when it may be possible to anticipate landslide-producing storms with sufficient confidence to alert agencies who can then inform those at risk to prepare. Such an approach is

* Correspondence to: W. E. Dietrich, Department of Earth and Planetary Science, University of California at Berkeley, 383 McCone Hall, Berkeley, CA 94720-4767, USA.

being done for flooding (e.g. Miller and Kim, 1997b; Vicente and Scofield, 1998), but the challenge for landslide hazard warning is much greater. Potential flood areas are easily mapped, and once delineated they do not tend to change through time as a consequence of subsequent flood events (unless changes in flood prevention infrastructure occur). Floods take time to build up because they are the result of the integration of runoff from widespread areas. In contrast, potential landslide locations are much harder to delineate because of the strong influence of local controls (soil thickness, root strength, localized seepage forces and bedding or fractures) and their threshold dependency. Furthermore, the hydrologic processes that lead to landsliding are local, and the time scales of response to storm precipitation variations may be just minutes.

Nonetheless, there are practical and scientific reasons to attempt to build a dynamic, storm-driven landslide model. Population pressure is leading to the expansion of development into landslide-producing environment. This is leading to conflict, for example over land use where houses abut upland areas used for timber production or other agricultural purposes. In part, owing to the great strides made in atmospheric sciences in recent years, the public has become accepting of the idea that the intensity of seasonal weather, such as El Niño cycles, has an element of predictability that permits advanced preparation of its consequences (Piechota *et al*., 1997; Delecluse, 1999). The scientific opportunity here is to explore further what controls the temporal and spatial distribution of landsliding across a landscape. Progress in understanding these controls will have clear practical application, but will also increase our ability to model landscape evolution driven by stochastic rainfall (e.g. Tucker and Slingerland, 1997).

Three general approaches for predicting landslide-producing storms have been taken: (1) empirical analysis of landslide-producing storm characteristics; (2) empirical mapping of landslide locations; and (3) mechanistic modelling of hydrology and slope stability. Here, we review each approach. The analysis of the relationship between landsliding and precipitation has suggested the use of rainfall thresholds to forecast the timing of abundant debris flow occurrence on a regional scale. Such an approach does not explicitly account for soil mechanics, and relies instead on historical records of past storms and landslides. The role of antecedent rainfall was emphasized by Campbell (1975), who observed that in southern California 10 in (254 mm) of antecedent seasonal precipitation were needed in order to bring the colluvium mantle to field capacity, so that subsequent storms may trigger landslide occurrence. Wieczorek (1987) observed a seasonal threshold of 280 mm at the La Honda test site in northern California. Canuti *et al*. (1985) developed an index that accounts for the weighted cumulative rainfall in the last 15 days, where the most recent data are given a higher weight. A similar approach was adopted by Crozier (1999) and Glade *et al*. (2000) in their Antecedent Soil Water Status Model (ASWSM), which accounts for the draining of early rainfall and accumulation of late precipitation. They propose an equation to estimate the probability of landsliding as a function of daily intensity and previous water accumulation. Crozier (1999) used this method over a single season (January–August 1996) and was able to identify the period with over ten documented landslides per day, although the observed occurrence of a number of isolated landslides was not predicted. Glade *et al*. (2000) applied the ASWSM over three different sites in New Zealand, performing a separate best-fit calibration for each area. Comparison of the calibrated thresholds for the three different areas suggests that the values are only of local significance. The selection of the number of days relevant to an antecedent rainfall may vary with the thickness of the soil mantle to be brought near saturation (Terlien, 1997) and its hydraulic conductivity (e.g. Iverson, 2000).

Caine (1980) collected a worldwide inventory of landslides and plotted the related rainfall conditions over an intensity–duration (*I–D*) graph, finding a threshold envelope able to discriminate critical triggering conditions for soil slips. He argued that such a relationship is reliable for durations shorter than 10 days and longer than 10 min, because (a) antecedent precipitation is not considered and (b) the intensity used is the average over the entire storm, not the peak intensity on a sub-hourly basis, so its calculation is strongly biased for short or long duration events.

Following the same approach, other local indexes have been proposed in order to obtain a better fit with the regional inventories, rather than using Caine's global data set (e.g. Nilsen and Turner, 1975; Cancelli and Nova, 1985; Larsen and Simon, 1993). Cannon and Ellen (1985) proposed a threshold for 'abundant' debris flow occurrences in northern California that was adopted as a basis for a warning system in California. They used the Campbell (1975) approach to discard any storm occurring before the required antecedent seasonal rainfall, and a calibrated local *I–D* curve for the remaining storms. The rainfall intensity values were normalized by the

mean annual precipitation (MAP) in order to account for different climatic domains (e.g. northern and southern California). Wieczorek (1987) analysed 7 years of rainfall at a test site and found that landslide-producing storm prediction was improved by considering the intensity that is exceeded during the entire duration used in the *I–D* plot. He proposed a local threshold for storms that might trigger as little as one debris flow. Some have found (Church and Miles, 1987; Kim *et al*., 1992) that specific areas may be more susceptible to landsliding because of high antecedent rainfall whereas other areas respond to short-term rainfall characteristics regardless of previous precipitation. The coupled antecedent rainfall intensity–duration threshold approach was used as a basis of a warning system in the San Francisco Bay area, where an alarm was successfully issued by the US Geological Survey in February 1986 (Keefer *et al*., 1987). Nevertheless, some limitations of this approach have to be considered. The empirical nature of these rainfall threshold approaches limits their general applicability. Wilson (1997, 2000) has proposed a procedure for the transfer to southern California of critical rainfall thresholds calibrated in northern California, taking into account the difference in MAP and annual number of rainy days (RDN) in order to explain the different pattern of debris flow response to different climates. Based on this rationale, a California map of rainfall thresholds for debris flow triggering has been published (Wilson and Jayko, 1997). It is not clear how well this approach will work in general, although such a procedure is necessary to generate regional thresholds where data are lacking. More importantly, this empirical approach may be applied to a particular region but not for any particular hillslope. Hence, if a warning is issued, because the spatial pattern of risk is unclear, it is difficult to implement the warning. Given the extremely high variability of risk associated with different topographic locations, without some map location of landslide potential, it would not be possible to to warn those individuals most at risk.

Several approaches have been proposed to map the spatial pattern of landslide potential, based on observed landslide occurrences (e.g. Carrara, 1983; Carrara *et al*., 1991, 1995; Van Westen, 1993; Chung *et al*., 1995). These methods produce landslide hazard maps, but they do not try to predict the timing of the events. They are based on the observation of past landslide events, and, assuming the same conditions still hold, assign a given degree of danger to all hillslope elements that feature analogous conditions. By using geographical information system (GIS)-based map overlay techniques it is possible to combine several maps of different parameters (e.g. topographic gradient, land use, bedrock geology, distance from faults and stream, etc.), with mapped landslides locations, to produce 'unique condition' polygons or grid cells. A rank index, used as a proxy for landslide hazard, can be derived from combining an assigned index to each of the predictors, or by using statistical techniques (e.g. Carrara, 1983; Carrara *et al*., 1991, 1995; Van Westen, 1993; Chung *et al*., 1995).

The rapidly growing availability of relatively detailed digital elevation data and of computing power has led to advances in mechanistic modelling of shallow landslide hazard, through coupling simple mechanistic slope stability and hillslope hydrological models (e.g. Okimura and Kawatani, 1987; Dietrich *et al*., 1992, 1993, 1995; Hsu, 1994; Montgomery and Dietrich, 1994; Terlien *et al*., 1995; Wu and Sidle, 1995; Duan, 1996; Pack *et al*., 1998). For this purpose, the simplest assumption for hillslope hydrology is to treat the subsurface flow in the steady state, and map the topographic control on the pore pressure (Dietrich *et al*., 1992, 1993, 1995, 2001; Montgomery and Dietrich, 1994; Pack *et al*., 1998). This pore pressure is then used in the infinite slope equation to estimate slope stability and produce maps of relative potential of shallow landsliding (e.g. Montgomery *et al*., 2000; Dietrich *et al*., 2001). This approach permits uncalibrated predictions and has proven reasonably successful, though there is a tendency for overprediction to occur, depending on the quality of topographic data (e.g. Dietrich *et al*., 2001).

Dynamic hydrological models have been proposed that predict landslide potential for specific rainstorms. Okimura and Kawatani (1987) computed subsurface runoff using a finite-difference grid scheme for slope-parallel, saturated subsurface flow. Rainfall recharge was routed instantaneously to the saturated zone, neglecting evapotranspiration and storage in the unsaturated zone. Assuming constant soil thickness, sites were classified within 'degree of danger' categories according to the respective time needed to reach instability given a constant rainfall of 20 mm h^{-1}. In later applications, Okimura (1989) explicitly measured soil thickness in the field because of its crucial role in the slope stability computations. Terlien *et al*. (1995) included Okimura's soil water balance model into a GIS-based module for mapping the probability of failure based on a regional database of geotechnical soil parameters. They noted that a scaling factor had to be applied to the hydrological variables of

the model (namely the storage capacity and the fraction of effective recharge reaching the saturated zone) in order to obtain a good fit with observed data. Wu (1993) and Wu and Sidle (1995) used the TAPES-C (Moore *et al*., 1988) contour-based scheme to calculate slope stability in forested catchments, incorporating the effects of forest dynamics on root strength contribution. Time-dependent variations of root strength associated with clearcut logging were modelled using a root decay model by Sidle (1992). Rainfall recharge was routed instantaneously to the saturated zone, neglecting evapotranspiration, unsaturated flow and storage in the unsaturated zone. Subsurface saturated flow was routed using a kinematic model based on Darcy's law. Minor spatial variations in soil depth and hydraulic conductivity were estimated from soil maps. They applied this model to a site in Oregon over a single landslide-producing storm (21 h), although the exact location of the slide could not be observed. Hsu (1994) developed a hydrological model that explicitly routed saturated subsurface flow over two layers (colluvium and bedrock) through a grid system, neglecting evapotranspiration and unsaturated storage. Unsaturated flow travel time was accounted for, but the discrepancies between theoretical and observed time to peak led to the incorporation of an empirical constant to accelerate the calculated response. Soil depth was calculated using the Dietrich *et al*. (1995) topographically driven transport model. The hydrological model was calibrated with the observed time series of mapped saturated areas, with some success, whereas the landslide model was tested with two sets of landslides mapped for different storms at the same catchment. Hsu (1994) did not test to see if her model successfully predicted failure occurrences over more than one storm using the same parameter set. She concluded that a simpler, steady-state model performed as well as a dynamic model, while having the advantage of needing fewer parameters. Duan (1996) developed a grid-based model that simulates surface and subsurface hydrology and snowmelt to predict pore pressure driving shallow slope stability. Unsaturated recharge rate was approximated using the Brooks and Corey (1964) relationship for moisture-dependent hydraulic conductivity. Soil depth and properties were deduced from SCS soil maps. Model calibration was performed using daily streamflow discharge data. Duan (1996) found that the most important parameters in his model are soil depth, hydraulic conductivity and its decline with depth, as well as soil and root cohesion. Friction angle and density of colluvium are also important, and their ranges are easier to determine or estimate. Porosity and field capacity were the less critical parameters. The model was not tested over a time series of precipitation versus landslide occurrence to check the performance of the model prediction.

Although the above authors considered only the pore pressure deriving uniquely from the build up of a saturated layer above a predefined critical slip surface, others (e.g. Crosta and Marchetti, 1993; Pradel and Raad, 1993; Rahardjo *et al*., 1995; Iverson, 2000) preferred to model the development of pore pressure as generated by the advance of a wetting front. Pradel and Raad (1993) used the Green–Ampt infiltration scheme to predict the critical timing of rainfall needed for the wetting front to reach the critical depth for triggering failure. Based on such theory, Crosta (1998) parameterized different rainfall thresholds over different areas based on the variability of soil properties. The role of transient unsaturated–saturated flow was considered explicitly by Rahardjo *et al*. (1995), who included negative pore pressure and its change with depth in the infinite slope equation. Their model predicts failure whose depth is not necessarily predefined by a discontinuity but instead is related to the intensity and duration of rainfall. Iverson (2000) has proposed a theory to predict rainfall-induced pore pressure in shallow soils by solving Richard's equation under the assumption of tension-saturated soils. Under this assumption, the model predicts the magnitude and timing of the pore pressure response to rainstorm, using only three parameters (rainfall intensity, duration and soil hydraulic diffusivity). The assumption of tension saturation narrows the applicability of Iverson's model, as first it would be necessary to predict which storms lead to tension saturation. Furthermore, although the incorporation of pore pressure response to transient unsaturated flow enables the effect of high-intensity short-duration rainfall to be estimated, it requires at least hourly precipitation data, which are not available for most sites, where calibration could be performed over past debris flow events.

Model performance will be strongly influenced by the quality of the data used for calibration. Landslide inventories are used by almost every landslide hazard zonation procedure, whether statistical (as direct input to a model) or deterministic (as calibration or validation data sets). Often, the quality of such data varies spatially, in that landslide mapping is most accurate for areas surrounding urban areas or economically important settlements, where risk perception and accuracy of mapping are greater (Glade, 1997). Moreover, the

resolution of the data collection will influence the inventory. Mapping from aerial photographs, the most popular data acquisition technique, may hinder the recognition of failures covered by thick vegetation, and the delineation of their boundaries may vary dramatically from operator to operator (Carrara, 1993; Carrara *et al*., 1995).

Annual landslide inventories are rarely done; instead, maps are generated after major landslide events, sometimes years later. This means that there is a poor record of the relationships between landslide locations and frequency of failure versus the precipitation record. Moreover, the underlying assumption for the use of such data in order to predict future failures is that similar conditions to those acting at the timing of past failures apply to future events. Land use, vegetation dynamics and effect of previous landsliding may significantly alter relative stability. Finally, the resolution of the data used as input or predictors for the varous models is crucial. The spatial and temporal pattern of rainfall intensity can vary substantially within a few kilometres, adding further difficulties in extrapolating data from sparse raingauge stations. The computation of topographic variables is extremely sensitive to the detail of the digital terrain model (DTM) and to the procedure used to extract them. For these reasons, efforts have been made towards refining topographic data capture in order to reduce DTM-derived uncertainty (Dietrich *et al*., 2001).

Here, we explore the possibility of using a coupled hydrologic–slope stability model to predict the time series of landsliding at a given ungauged catchment, where three landslide events were mapped throughout a 50 year record. Although single landsliding episodes have been compared before with dynamic distributed modelling (Wu, 1993; Duan, 1996), this appears to be the first exercise dealing with repeated events within the same catchment. This is a difficult challenge because: (a) large uncertainty is present in parameter estimation for a coupled hydrology–slope stability model; (b) soil properties and land use may change with time; (c) the accuracy of scars recognition and mapping is limited by the quality of topographic maps used and by the visibility from aerial photographs; and (d) the quality of topographic and rainfall data for past years may be poor.

The innovation of using a time series at one catchment enables us to tackle the problem of 'false positives', i.e. prediction of landslides when none occurred, and to evaluate whether calibration on a single landslide event gives reliable predictions in future storms. Through extensive calibration, our model performs better than empirical *I–D* threshold models. We conclude, however, that poor data on topography, soil properties, precipitation, hydrological properties and changing site conditions through time will prevent models with dynamic hydrology from producing accurate predictions of the timing and location of shallow landslides. For practical purposes, therefore, simple rainfall thresholds coupled with simple models that predict the areas most likely to fail may prove more useful.

THE MODEL

The model used here couples the infinite slope stability equation with a dynamic hydrological model. It represents an attempt to model the hydrological response of hillslopes with a minimum set of parameters, using widely available data, namely 5–10 m grid digital elevation models (DEMS) and daily precipitation records. The hydrological model follows the two-layer conceptual scheme used by Hsu (1994), but uses an organization of storage similar to TOPMODEL (e.g. Beven and Kirkby, 1979; Beven, 2001), accounting for evapotranspiration and unsaturated zone storage. Both the hydrologic and slope stability models require estimates of local soil depth. Here, we predict the soil depth using the soil production and transport model of Dietrich *et al*. (1995), subsequently validated by Heimsath *et al*. (1999). Local piezometric response is calculated by a distributed rather than lumped model (e.g. Lamb *et al*., 1997, 1998); as Wigmosta and Lettenmaier (1999) pointed out, an explicit routing scheme represents more faithfully the water table fluctuations, and allows a better calibration of hydrological parameters.

We use daily rainfall data because of the widespread availability for the analysis of historical events. It is known from field and laboratory observations (e.g. Johnson and Sitar, 1990; Marui *et al*., 1993; Reid *et al*., 1997; Torres *et al*., 1998) and theory (e.g. Iverson, 2000) that pore pressure may respond very rapidly to intense and short (sub-hourly) rainfall pulses. Such rainfall data are rarely available, and it is later suggested that the absence of accurate high-resolution data on such controls as topography and geomechanical properties negates the practical value of a model capable of routing short rainfall pulses.

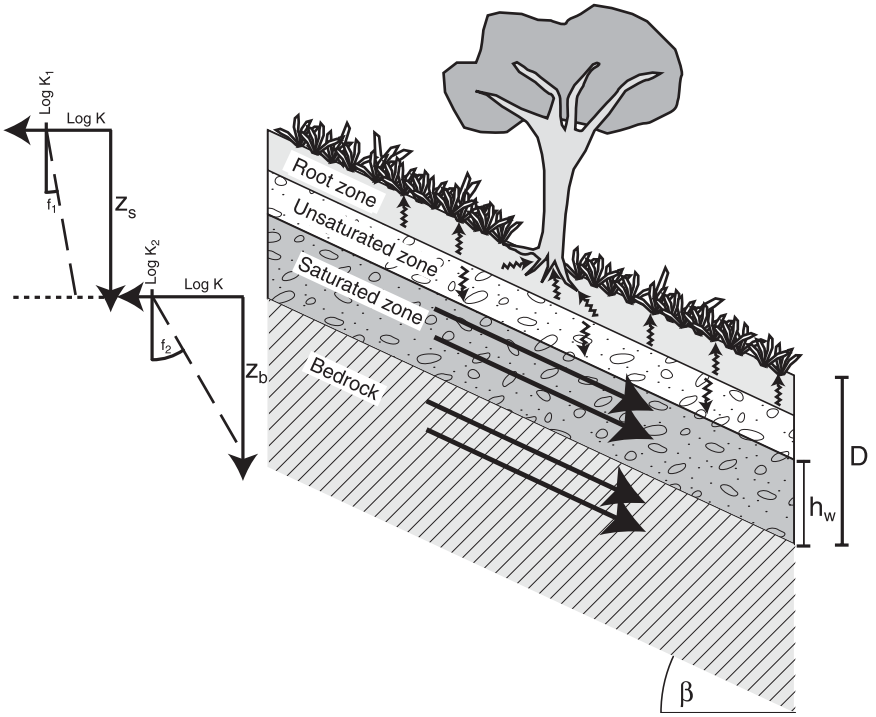

Figure 1. Elements of the coupled hydrological–slope stability model. The hydrological model accounts for evapotranspiration, unsaturated storage and routes slope–parallel subsurface discharge in soil and bedrock, calculating water table height over the soil–bedrock interface. This value is translated into pore-pressure in the infinite slope equation to calculate slope stability

Slope stability

The infinite slope model (e.g. Skempton and DeLory, 1957) calculates the stability of a soil prism (Figure 1) characterized by friction angle ϕ', bulk density ρ_s, and cohesion C'. The slip surface is coincident with the soil–bedrock interface. Subsurface water flow is assumed to be parallel to the ground surface. The factor of safety FS gives the ratio between resisting and driving forces:

$$FS = \frac{C' + (D\rho_s - h_w\rho_w)g\cos^2\beta\tan\phi'}{\rho_s g D\sin\beta\cos\beta} \quad (1)$$

where C' (N m^{-2}) is the effective cohesion, g (9·81 m s^{-2}) is the gravitational acceleration, ρ_s and ρ_w (kg m^{-3}) are the bulk densities of soil and water respectively, D (m) is the soil depth, h_w (m) is the water table height above the slip surface, β is the inclination of the topographic surface, the subsurface flow lines and the slip surface, and ϕ' is the effective friction angle. Equation (1) is widely used to analyse the stability of shallow soils using DTMs (e.g. Hsu, 1994; Montgomery and Dietrich, 1994; Dietrich *et al*., 1995; Terlien *et al*., 1995; Wu and Sidle, 1995; Duan, 1996; Pack *et al*., 1998; Iverson, 2000).

The time-dependent part is expressed in terms of the water table level h_w, which fluctuates seasonally according to precipitation. Whenever FS < 1, the driving forces prevail and failure occurs. The infinite slope model has been widely employed to compute FS of shallow landslides in colluvium and in bedrock, provided the ratio between failure length and depth is high enough (≫10).

The model assumptions (slip surface and flow lines parallel to the ground surface, null interslice forces) often do not hold. The model is not applicable to deep-seated landslides, whose slip surface geometry may differ substantially from a planar surface. From the hydrological point of view, flow field variations may cause pore pressure increases that a simple slope-parallel flow model is unable to compute (Iverson, 1990). This is

corroborated by field experiences (Wilson, 1988; Torres *et al*., 1998) that showed that exfiltration from bedrock may be the primary factor of unexpected shallow failures. Moreover, this model does not take into account the three-dimensional geometry of the landslide, neglecting lateral forces. Most important for shallow landslides modelling in coarse colluvium is that root strength dominates the cohesion term, and where the soil is sufficiently thick it acts primarily along the scar sides rather than along the slope-parallel sliding surface, as assumed in Equation (1) (e.g. Reneau and Dietrich, 1987; Schmidt *et al*., 2001). In order to minimize the number of parameters in the model, we elect not to address these issues here.

Hydrology

The model developed in this work adopts the storage organization (Figures 1 and 2) commonly used by many TOPMODEL applications (e.g. Beven, 2001). This consists of: a *Root Zone*, where evapotranspiration occurs; an *Unsaturated Zone*, located between ground surface and the water table; a *Saturated Zone*, bounded by the piezometric surface. We assume that there is always a saturated zone, located above or below the soil–bedrock interface. Horton overland flow is not considered in this application, and runoff resulting from saturated overland flow is simply removed instantaneously from the catchment, and ignored in further computations.

When the rainfall infiltrates the ground surface, it is stored into the *Root Zone*. This reservoir has a maximum capacity S_{rm}, and whenever the actual water content S_r exceeds the maximum capacity, the excess is transferred directly to the *Unsaturated Zone*. Such flow increases the water content S_u of the *Unsaturated Zone*. The flow rate Q_u in this zone is assumed vertical. As S_u approaches the maximum unsaturated storage S_{um}, hydraulic conductivity approaches the saturated hydraulic conductivity. At each site, maximum unsaturated storage S_{um} is given by the product of soil depth D and effective porosity n_{wet}. The *Saturated Zone* receives water from the above *Unsaturated Zone* (Q_u) and from the lateral transport from upslope neighbours, and in turn delivers water to the downslope *Saturated Zone*. Saturated flow is computed through Darcy's law. The model uses the same distinction adopted by Hsu (1994) and Dietrich *et al*. (1995), where a two-layer model was adopted to describe subsurface flow, assuming exponential decay of hydraulic conductivity with depth in the soil and in the bedrock layer, as suggested by field observations (e.g. Wilson, 1988; Montgomery, 1991; Torres *et al*., 1998). The importance of bedrock flow for the initiation of shallow landslides has been observed and documented by several authors (e.g. Wilson, 1988; Montgomery, 1991; Dietrich *et al*., 1995; Torres *et al*., 1998). Evapotranspiration has been neglected by some event-based hydrological models (e.g. Hsu, 1994; Wu and Sidle, 1995), on the grounds that, during a single storm event, it is unlikely to be a major controlling factor. Because the model we used is intended to work on a continuous record of years, however, the effect of evapotranspiration is included here. We used the most economical approximation of evapotranspiration, an empirical formula proposed by Hamon (1961), and used, amongst others, by Wolock (1993) in a TOPMODEL implementation. Other, more precise, methods can be used (e.g. Penman, 1948; Monteith, 1965), at the cost adding more parameters and uncertainty to the model. Hamon's (1961) approximation yields a rough daily estimate of potential evapotranspiration given the least number of parameters (daily temperature and latitude of the site). Following Haith and Shoemaker (1987), the Hamon (1961) estimate of potential evapotranspiration is

$$E_p = \frac{0.0021 H_t^2 e_s}{(T_t + 273.2)} \tag{2}$$

where E_p (m day^{-1}) is the potential evapotranspiration, T_t (°C) is the temperature at day t, H_t is the average number of daylight hours per day during the month in which day t falls, and e_s (kPa) is the saturated vapour pressure at temperature T. H_t can be calculated by using the maximum number of daylight hours N_t:

$$N_t = \frac{24\omega_s}{\pi} \tag{3}$$

where ω_s is the sunset hour angle of day t calculated as

$$\omega_s = \arccos(-\tan\phi \tan\delta) \tag{4}$$

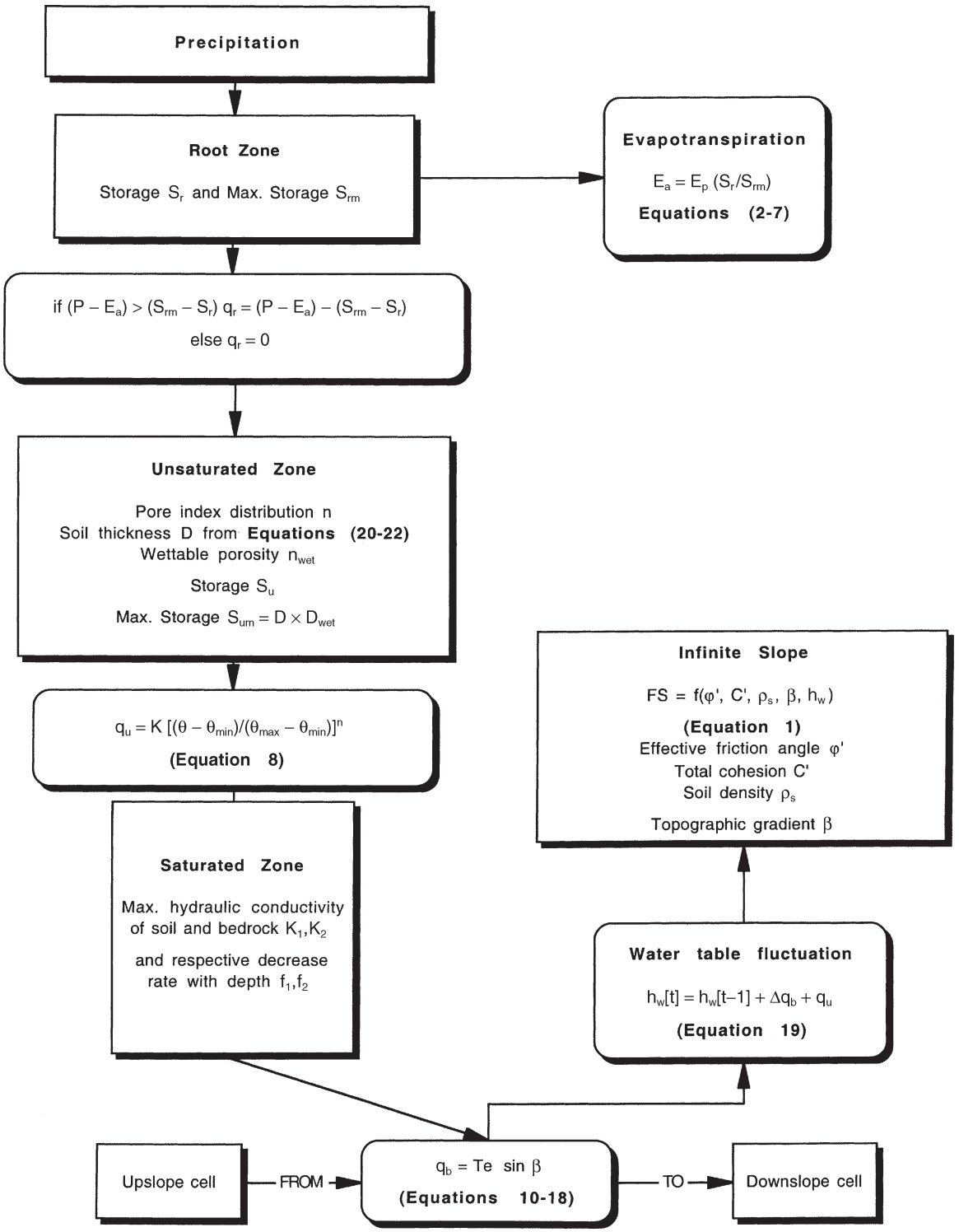

Figure 2. Flow diagram showing the steps used by the model in hydrological and slope stability calculation. The terms are defined in the text

ϕ is the latitude of the site and δ is the solar declination on day J (Julian day) of the year, expressed as

$$\delta = 0{\cdot}4093 \sin\left(\frac{2\pi}{365} J - 1.405\right) \tag{5}$$

e_s is computed as a function of temperature T:

$$e_s(T) = 0{\cdot}6108\, e^{(17{\cdot}27T/273{\cdot}3+T)} \tag{6}$$

Thus, given the latitude of the site and the average temperature, it is possible to formulate a rough estimate of potential evapotranspiration.

Actual evapotranspiraton E_a (m day^{-1}) is considered only within the root zone, and is computed as a fraction of potential evapotranspiration given by Equation (2), according to the ratio of actual versus maximum storage in the root zone (e.g. Beven, 2001).

$$E_a = E_p \frac{S_r}{S_{rm}} \tag{7}$$

where S_r and S_{rm} (m) are the actual and maximum water storage in the root zone. When water content exceeds the maximum storage in the root zone, the excess q_r is transferred to the unsaturated zone, where evapotranspiration is neglected.

Unsaturated flux per unit surface q_u (m day^{-1}) is assumed vertical and is approximated using a simple Brooks and Corey (1964) type equation:

$$q_u = K\left(\frac{\theta - \theta_{min}}{\theta_{max} - \theta_{min}}\right)^n \tag{8}$$

where K (m day^{-1}) is the saturated hydraulic conductivity, θ, θ_{min} and θ_{max} (m) are respectively the actual, minimum and maximum soil water contents in the unsaturated zone, and n is the pore distribution coefficient dependent on the spatial arrangement of grains, variable between 3·7 and 5·5 for most loamy/sandy soils according to Parlange *et al*. (1999). This simplified equation has been used by Grayson *et al*. (1992).

The term $(\theta - \theta_{min})$ appearing in Equation (8) is the actual unsaturated storage S_u, whereas $(\theta_{max} - \theta_{min})$ is the maximum unsaturated storage S_{um}. When the soil column is saturated, $S_{um} = S_u$ and unsaturated flux equals saturated flux. No time lag is accounted for in the travel time between ground surface and water table. The same scheme is adopted by, among others, Wolock (1993) and Beven (2001). Total vertical recharge rate to the saturated zone Q_u (m^3 day^{-1}) is then given by

$$Q_u = q_u A \tag{9}$$

where A (m^2) is the area of the elementary spatial unit, in this case a grid cell.

Saturated flow is computed according to Darcy's law, assuming slope-parallel flow, when subsurface discharge per unit width q_b (m^2 day^{-1}) is given by

$$q_b = T_e \sin\beta \tag{10}$$

where β is the hydraulic gradient. T_e (m^2 day^{-1}) represents the integration of hydraulic saturated conductivity over the entire saturated thickness.

$$T_e = \int_{\infty}^{D} K\, dz \tag{11}$$

where z (m) is depth measured vertically downward and D (m) is vertical soil depth and coincides with the depth of the critical soil slip. In a two-layer conceptual scheme (Figure 1), Equation (11) has to be calculated for both the soil mantle and the bedrock.

The model assumes exponential decay with depth of hydraulic conductivity for both the colluvium and the bedrock, at different rates.

$$K = K_1 e^{-f_1 z_s} \tag{12}$$

$$K = K_2 e^{-f_2 z_b} \tag{13}$$

In these equations, depth and hydraulic conductivities are measured separately for the two layers. K (m day^{-1}) is the saturated hydraulic conductivity of soil in Equation (12) and of bedrock in Equation (13); f_1 and f_2 (m^{-1}) are the decline rates of hydraulic conductivity in soil and bedrock respectively; z_s (m) is the depth measured from the ground surface within the soil layer, and z_b (m) is the depth measured from the bedrock–soil interface.

Combining Equation (11) with Equations (12) and (13) yields

$$T_e = \int_{h_1}^{D} K_1 e^{-f_1 z_s}\, dz + \int_{h_2}^{\infty} K_2 e^{-f_2 z_b}\, dz \tag{14}$$

where h_1 and h_2 (m) are the depths of water table from the ground surface and the soil–bedrock interface respectively.

Integrating Equation (14) yields

$$T_e = \frac{K_1}{f_1}(e^{-f_1 h_1} - e^{-f_1 D}) + \frac{K_2}{f_2}(e^{-f_2 h_2}) \tag{15}$$

If the water table is above the soil–bedrock interface, then $h_2 = 0$ and Equation (15) becomes

$$T_e = \frac{K_1}{f_1}(e^{-f_1 h_1} - e^{-f_1 D}) + \frac{K_2}{f_2} \tag{16}$$

If the water table is coincident or below the soil–bedrock interface, then $h_1 = D$ and Equation (15) becomes

$$T_e = \frac{K_2}{f_2} e^{-f_2 h_2} \tag{17}$$

Subsurface flow q_b can be calculated from Equation (10) and it is routed kinematically to transfer water downslope. Groundwater outflow discharge Q_b (m^3 day^{-1}) is computed as

$$Q_b = q_b b \tag{18}$$

where b (m) is the contour width of the flow element. At each time step, Q_b is delivered in proportion to the topographic gradient (Freeman, 1991; Quinn *et al*., 1991; Dietrich *et al*., 1995) to downslope neighbours, which therefore receive at the next time step some amount of subsurface inflow from upslope cells. Multiple flow direction apportioning has some shortcomings, such as dispersivity (Costa Cabral and Burges, 1994), but it reduces the sensitivity to grid orientation, compared with single flow direction algorithms (Freeman, 1991; Quinn *et al*., 1991; Dietrich *et al*., 1995). The outflow supply becomes available to the neighbours at the subsequent time step. At every time step t, water table elevation h_w (m) is computed as

$$h_w(t) = h_w(t-1) + \frac{\Delta Q_b(t) + Q_u(t)}{A} \tag{19}$$

where ΔQ_b is the difference between inflow from upslope cells and downslope outflow, and A is the area of the elementary unit. Equation (19) is then substituted into Equation (1) to calculate the factor of safety.

Soil thickness

The thickness of soil mantle, which in our model coincides with the failure depth, is a critical parameter in both slope stability and hydrological models (Hsu, 1994; Dietrich *et al*., 1995; Terlien *et al*., 1995; Wu and Sidle, 1995; Duan, 1996; Iverson, 2000). Soil thickness is rarely mapped, and it is a practical impossibility to do so over a large catchment. To estimate soil depth D we used the model proposed by Dietrich *et al*. (1995), which explicitly relates it to topographic curvature:

$$\rho_s \frac{\partial D}{\partial t} = -\rho_r \frac{\partial e}{\partial t} - \nabla \cdot \rho_s \tilde{q}_s \tag{20}$$

where ρ_s and ρ_r (kg m^{-3}) are the densities of colluvium and bedrock respectively; $\partial e/\partial t$ (cm year^{-1}) is the rate of soil production.

The simplest transport law for hillslope sediment transport $\tilde{q}_s$, adopted by Dietrich *et al*. (1995), assumes flux varies linearly with local slope ∇z:

$$\tilde{q}_s = -K_t \nabla z \tag{21}$$

where K_t (cm^2 year^{-1}) is a soil transport coefficient. Soil production is approximated as exponentially declining with depth H (m).

$$-\frac{\partial e}{\partial t} = P_0 \, e^{-\alpha D} \tag{22}$$

according to empirical findings by Heimsath *et al*. (1997, 1999), where P_0 (cm year^{-1}) is the maximum (surface) soil production and α represents exponential decline with depth.

An initial soil depth of 30 cm was assumed and the model was run for 15 000 years, as in Dietrich *et al*. (1995). The parameter values of the soil model were estimated within the range found by Heimsath *et al*. (1997, 1999): $P_0 = 0{\cdot}02$ cm year^{-1}; $\alpha = 0{\cdot}05$ m^{-1}; $K_t = 0{\cdot}019$ cm^2 year^{-1}.

APPLICATION OF THE MODEL

Site description

We calibrated and tested the coupled slope stability–hydrological model over a $1{\cdot}7$ km^2 wide catchment located near the city of El Granada, in the Montara Mountains (San Mateo County, California; Figure 3). Two debris flow events were mapped by Wentworth (1986) from aerial photographic interpretation combined with some field mapping (Figure 4). The first occurred during the 15–28 December 1955 storm, which caused numerous landslides over the entire Pacific Coast, reaching its peak in the Bay area on 22–23 December (Division of Water Resources, 1956). The second event was caused by the 3–5 January 1982 storm that caused in the San Francisco Bay area US$66 million in damages and the death of 25 people (Ellen and Wieczorek, 1988). It was highly localized and brief compared with the 1955 event. In addition, we observed a time series of detailed aerial photographs (1 : 24 000) from 1953 though 1998, which show that: (a) in 1955, 80% of the hollows failed; (b) no previous scars were evident from the 1953 aerial photographs, an observation confirmed by two aerial photographs taken in 1931 and 1946 reported as illustrations in Van der Werf (1992); (c) a third landsliding episode was visible from the aerial photograph taken in September 1998, which we estimate to be caused by the storm of 31 January 1997–3 February 1998. No other landslides appear to have occurred within the time window covered by the aerial photographs (1953–98). Many of the 1955 scars still look fresh, unrecovered, over the entire time series of the photographs.

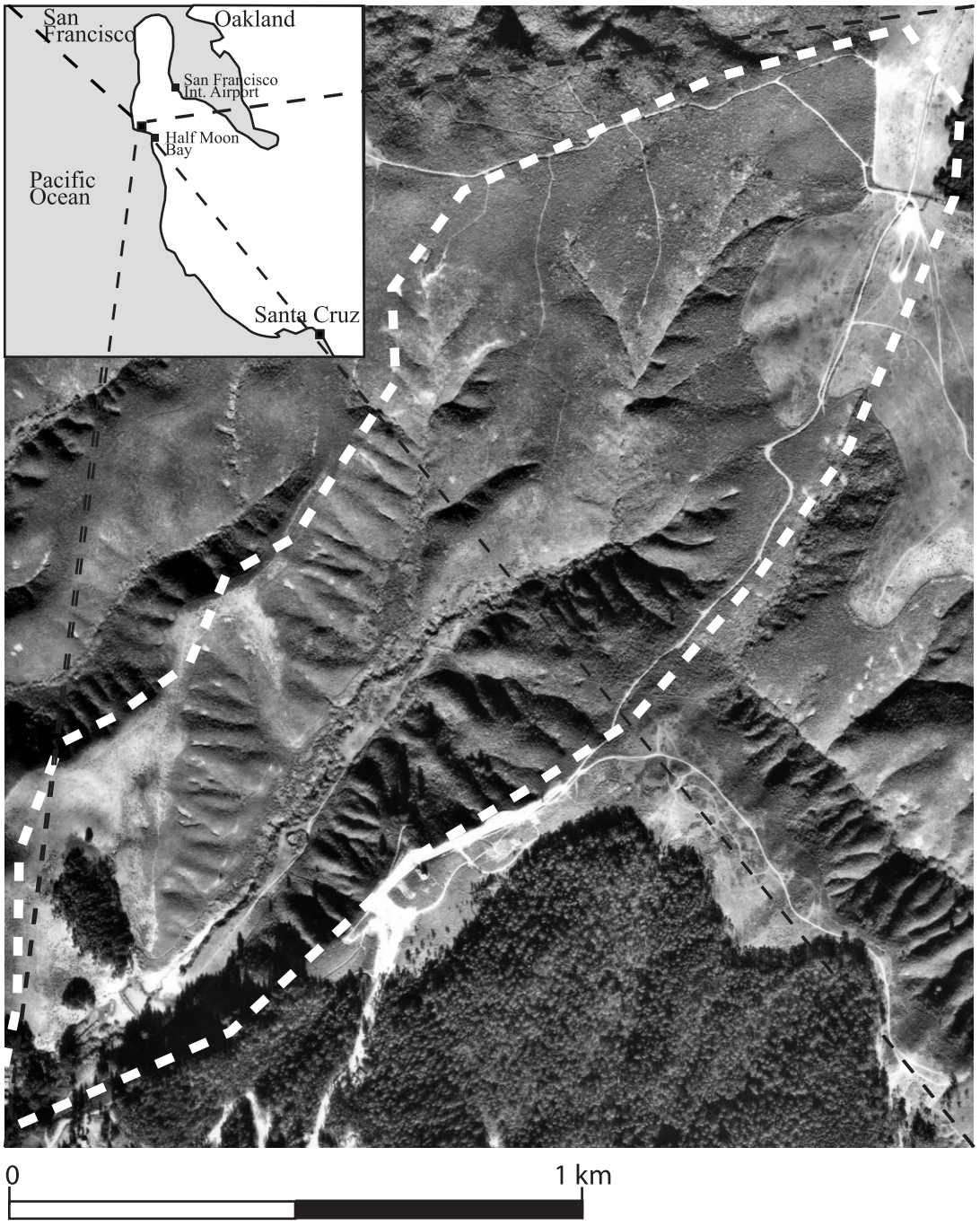

Figure 3. Location map of the test area located near El Granada (San Mateo County, California)

Bedrock lithology mainly consists of Montara Diorite, a medium to coarse granitic rock, composed of quartz diorite with some granodiorite, apatite and pegmatite (Wentworth, 1986). The rock is pervasively jointed and fractured and the topmost 3 to 30 m is moderately weathered to locally disintegrated in granular soil (Wentworth, 1986). The soil mainly comprises Miramar coarse sandy loam (Wagner and Nelson, 1961). Estimated hydraulic conductivities of the surface soil are classified by Wagner and Nelson (1961) as 'rapid and relatively low' which,

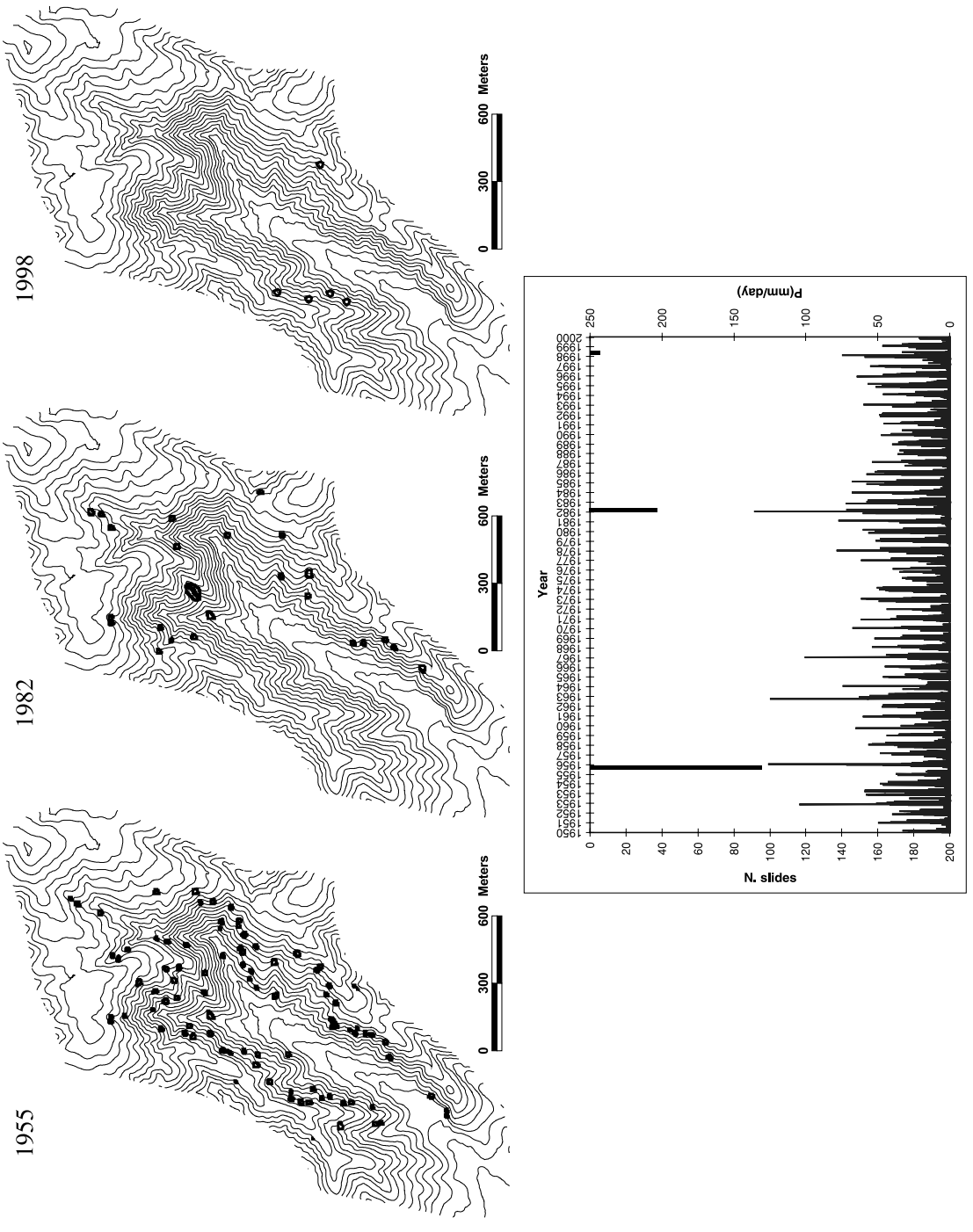

Figure 4. Landslide locations mapped for the three events by Wentworth (1986) for 1955 and 1982 events, and from aerial photographs for the 1998 event. Lower graph shows the daily precipitation from 1950 to 2000 and the number of landslides in the three events. The magnitude of landslide response does not seem related to the severity of the rainfall input. Contour elevation spacing: 10 m

according to the soil survey manual, are comparable to $(3{\cdot}5–7{\cdot}1) \times 10^{-5}$ m s^{-1} and $(1{\cdot}4–5{\cdot}6) \times 10^{-6}$ m s^{-1}. No data are available on the distribution of soil thickness.

Current vegetation consists mostly of shrubs and grasslands, with isolated patches of eucalyptus in the flatlands that extend from the fan area. According to Van der Werf (1992), before 1769, seasonal burning by Native Americans of native hillside vegetation limited the growth to bunch grasses and fire-resistant shrubs. Between 1769 and 1915 grazing was introduced by the Spanish and this was continued mostly by Mexican communities and the early American settlers. Grazing was limited in 1910–20, when the city of El Granada filed several complaints against local citizens for uncontrolled grazing. Since then, the vegetation of these hillsides has remained in its natural state.

The climate of the region is Mediterranean, with dry mild summers and moist cold winters. Average daily temperature ranges between 10 and 15 °C. Mean annual rainfall is 750 mm, 80% of which occurs between November and March (Wagner and Nelson, 1961). A preliminary comparison of the rainfall record versus the landslide history suggests a discrepancy between the rainfall and landslide time series, as there is poor correspondence between the most important debris flow pulse and the severity of rainfall. In order to explore this issue, we performed a rainfall analysis.

Rainfall analysis

We analysed the rainfall *I–D* pattern of the available rainfall records and compared it with the landslide-producing threshold curves proposed by various research groups. The closest daily record is at the Half Moon Bay station and the closest hourly record is from the San Francisco International Airport station (Figure 3). We calculated *I–D* envelopes using the procedure outlined in Chow *et al*. (1988), (e.g. intensity is averaged over a given duration time window). The daily *I–D* curves (Figure 5) were derived from the analysis of the entire hydrological year for every season. For durations longer than 1 day, the intensities of non-landslide-producing storms do not differ significantly from the 1982 storm event. Hence, we cannot derive any significant envelope for discriminating events likely to trigger debris flows. The *I–D* curves calculated on a hourly basis for the most severe storms between 1955 and 2000 (Figure 6) are compared with the thresholds suggested by Caine (1980), Cannon and Ellen (1985), Wieczorek (1987) (the latter two as reported in Keefer *et al*. (1987) and Wilson and Jayko (1997).

Only the 1982 event stands out relative to the rest. Generally, the 1955 storm is high, but not in exceededance of some other storms that did not provoke landslides. The 1998 storm was exceeded by many previous non-slide-producing events. Surprisingly, the best performing empirical threshold curve was the one proposed by Caine (1980), the only one of the above curves that was not based on local data. Nonetheless, the Caine curve model incorrectly predicts three events that did not produce landslides and misses two that did. None of the other curves performed well. This failure may be in part due to the inapplicability of the precipitation data to the local catchment where landslides occurred. Below, we explore whether a mechanistic model might be more successful in identifying landslide-producing storms.

Model parameterization and calibration

The model was run using a 5 m spaced grid derived by Hsu (1994) from the US Geological Survey contour maps. We used as climatic record the daily time series from 1950 to 1999 of precipitation (Figure 4) and temperature (to estimate evapotranspiration) recorded at the Half Moon Bay station (available from the National Climatic Data Center). The isohietal maps provided by Monteverdi (1984) for several rainstorms suggest that these values can be reasonably used for the El Granada catchment without any transformation.

Table I summarizes the parameter values (and their sources) used in the model. Only four hydrological parameters were allowed to vary; all other parameters were fixed. Colluvium effective porosity (needed for converting saturated water content into piezometric head) was kept constant at $n = 0{\cdot}2$, as were the geotechnical parameters ($\rho_s = 1800$ kg m^{-3}, $C = 1000$ N m^{-2} and $\phi' = 40°$), based on Hsu (1994) and Dietrich *et al*. (1995). Maximum root storage S_{rm} was set to 0·005 m. The four hydrological parameters calibrated to optimize model performance were K_1, K_2, f_1 and f_2. These are the saturated conductivities of the soil and bedrock and their corresponding declines with depth.

A landslide was considered as correctly predicted if the model predicted any of its cells as unstable ($FS \leq 1$) (e.g. Montgomery and Dietrich, 1994; Dietrich *et al*., 1995, 2001). We define an optimal model as one that is

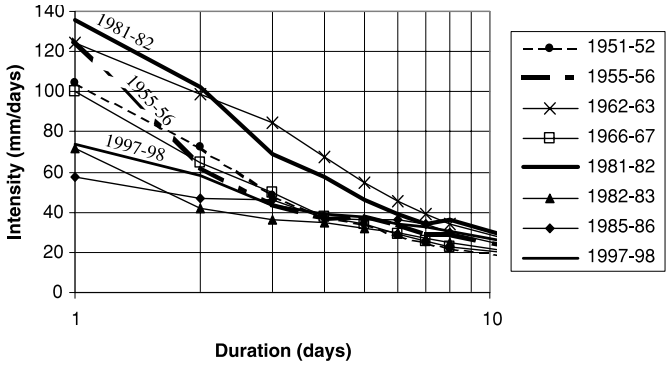

Figure 5. Intensity–duration graph based on daily rainfall data from the Half Moon Bay station. For the sake of clarity, minor storms were not plotted. The hydrological years featuring landslides are plotted with heavier lines. No clear envelope can be drawn that discriminates landslide-producing sequences from the others

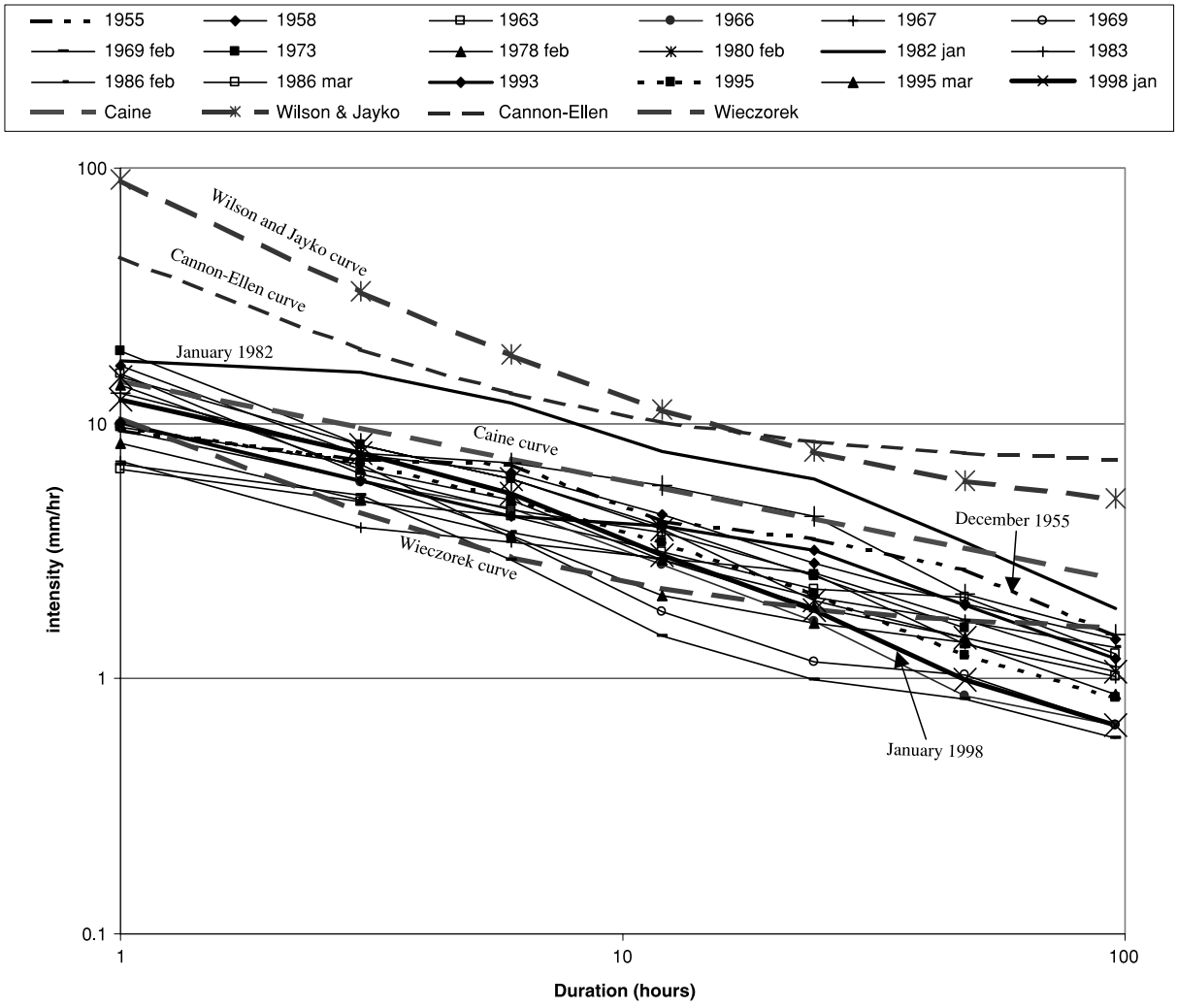

Figure 6. Intensity–duration graph based on hourly rainfall from the San Francisco International Airport station. The data are compared with Caine (1980), Cannon and Ellen (1985), Wieczorek (1987) and Wilson and Jayko (1997). The Caine curve discriminates correctly for 1982, whereas the others perform poorly

Table I. Parameter ranges used for model calibrations[a]

Process	Parameter	Value/range	Source
Rainfall	P (mm)	Variable	Half Moon Bay station (daily)
Topography	$\boldsymbol{\beta}$	Steepest descent	DEM from Hsu (1994)
	b (m)	5 m	DEM from Hsu (1994)
	$\boldsymbol{A}$ **(m²)**	Slope-proportional partitioning	DEM from Hsu (1994)
Evapotranspiration	T (°C)	Variable	Half Moon Bay station (daily)
Subsurface hydrology	K_1 (m day^{-1})	10–100	Calibrated
	K_2 (m day^{-1})	2–5	Calibrated
	f_1	0·5–20	Calibrated
	f_2	1·4–10	Calibrated
	n_{wet}	0·2	
	$\boldsymbol{D}$ **(m)**	Spatially variable	Soil depth module
	S_{rm}	0·005	
	S_{um}	Spatially variable	D^*n_{wet}
Soil depth	K_t (cm² year^{-1})	50	
	P_0 (cm² year^{-1})	0·019	Dietrich *et al.* (1995)
	α	0·05	
	t (year)	15 000	
Shear strength	C' (Pa)	1000	
	ρ_s (kg m^{-3})	1800	
	ϕ' (°)	40	

[a] Parameters in bold are distributed throughout the grid.

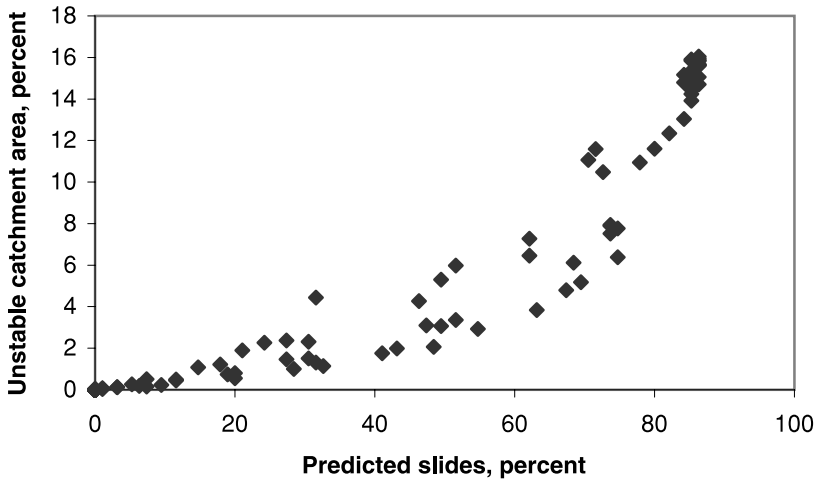

Figure 7. Performance in terms of predicted landslides and predicted unstable area in the overall catchment. The best performance was shown by 20 simulations, which were able to capture about 85% of the landslides while classifying as unstable about 15% of the catchment area

able to identify the maximum number of landslides with the minimal percentage of land predicted to fail. For instance, a model that predicts 100% of the mapped failures but which also classifies 80% of the catchment area as unstable obviously has no discriminant power.

The model consistently overpredicted slope instability, with some cells being classed as unstable every year. Two steps were taken to deal with this. First, based on field evidences of channel occurrence, we estimated that all cells draining areas greater than 1000 m^2 were channels. These cells were not used in the landslide analysis. Second, we used a cutoff that a storm has to predict more than 11% of the catchment to be unstable in order

to be considered a landslide-producing storm. We infer that the background 11% value is largely associated with poor topography and unknown local geomechanical and hydrological properties. The choice of 11% represents an arbitrary threshold, which may not be applicable elsewhere.

Calibration of the four hydrological parameters was first accomplished for the December 1955 event using 90 different combinations of K_1, K_2, f_1 and f_2. Those values were selected based on the values reported by Wagner and Nelson (1961) and used by Hsu (1994). Some 20 of these simulations were able to predict 84–86% of the actual mapped landslides successfull, while predicting 14–16% of the catchment as unstable (Figure 7). The date of the peak response calculated in such runs, however, varied by as much as 4 days. The results were sensitive to the ratios K_1/f_1, K_1/K_2 and f_2/f_1 (Figure 8). The ratio K_1/f_1 scales with the transmissivity of the colluvium layer

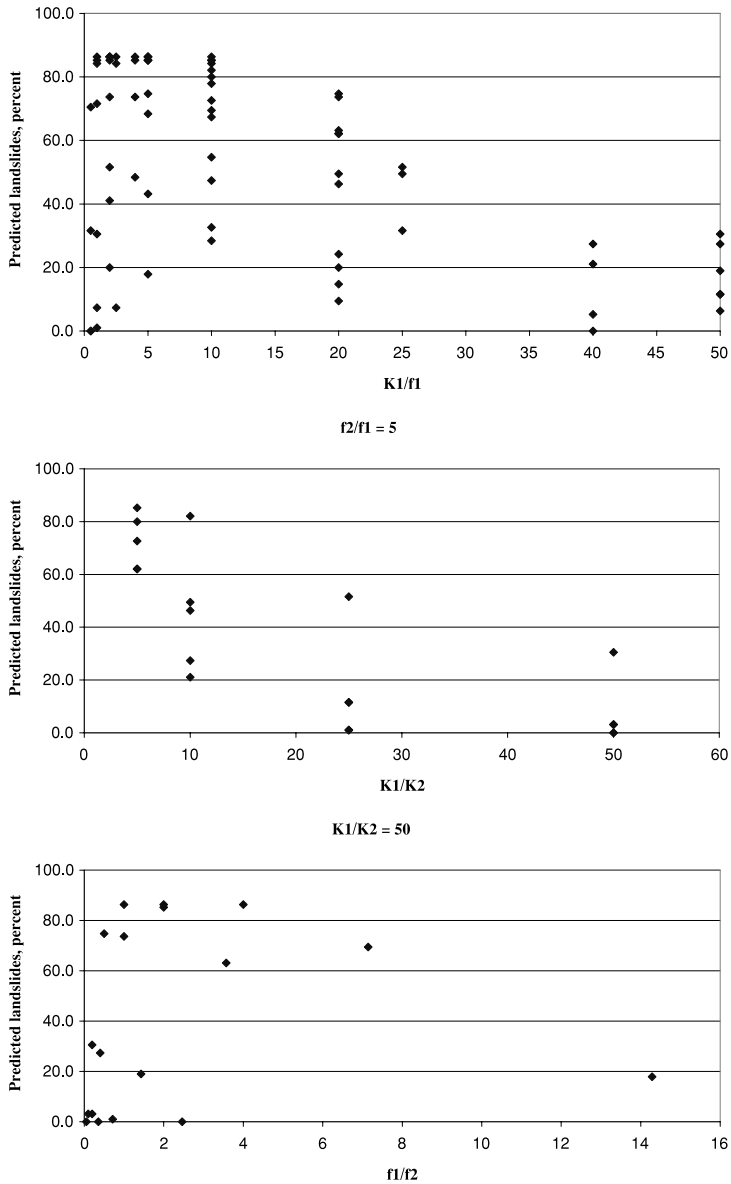

Figure 8. Effect of the ratios K_1/f_1, K_1/K_2 and f_2/f_1 on the predicted landslides. The K_1/K_2 and f_1/f_2 plots refer only to a subset of the simulation, keeping the other ratio constant. Best performance occurs for $K_1/f_1 < 10$ m day^{-1}. As the contrast in hydraulic conductivity K_1/K_2 decreases, the model performs better. Similar results are observed for the contrast in the rate of conductivity decline with depth, f_2/f_1

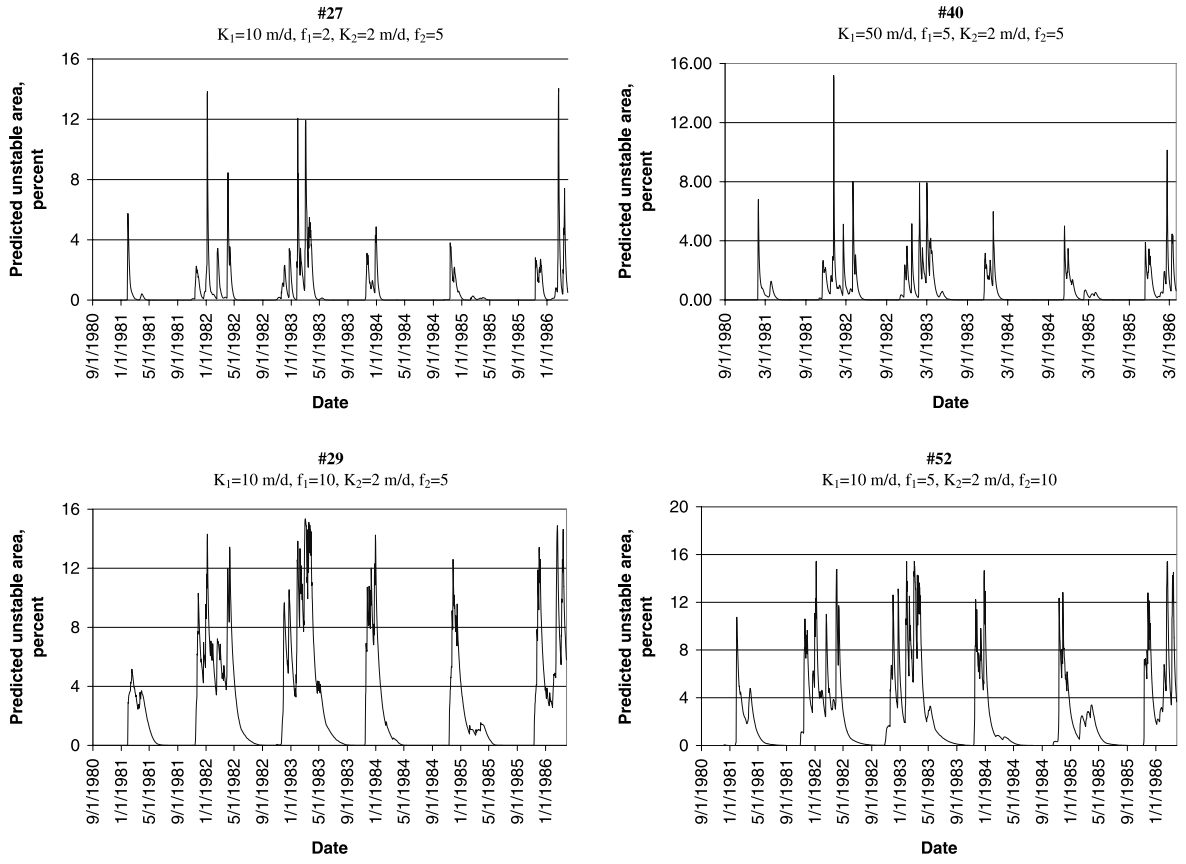

Figure 9. Examples of simulations run for the 1980–86 period, using the parameters set calibrated for the December 1955 landslide-triggering storm. Simulation no. 40 has the best predictive power and was selected as the best parameter set

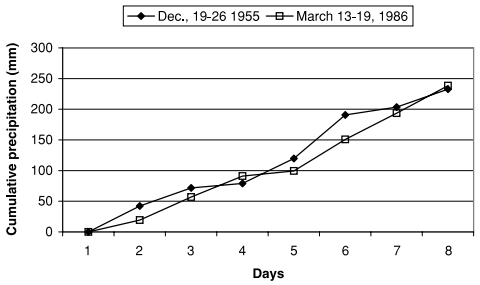

Figure 10. Comparison of the December 1955 and March 1986 storm characteristics. Despite the similarity, the 1986 storm did not trigger any debris flow in the test area

(actually, in a single-layer model, it is equal to the transmissivity). The simulations gave the best results for values of K_1/f_1 ranging between 2 and 10 m^2 day^{-1}, consistent with the values estimated from Wagner and Nelson (1961). The larger the differences between the soil saturated conductivity and bedrock conductivity, the poorer were the results. The same applies to the ratio f_2/f_1 between the rates of conductivity decline within the two layers.

These 20 parameter sets were then used to run 20 simulations (Figure 9) in another time window (1980–86) that contained three critical storms: 3–5 January 1982, 24 February–3 March 1983 and February 1986, of which only the first caused debris flows. In particular, the February 1986 storm was similar to the 1955 (Figure 10), while having an even greater antecedent cumulative seasonal rainfall (221 mm for 19 December 1955 versus

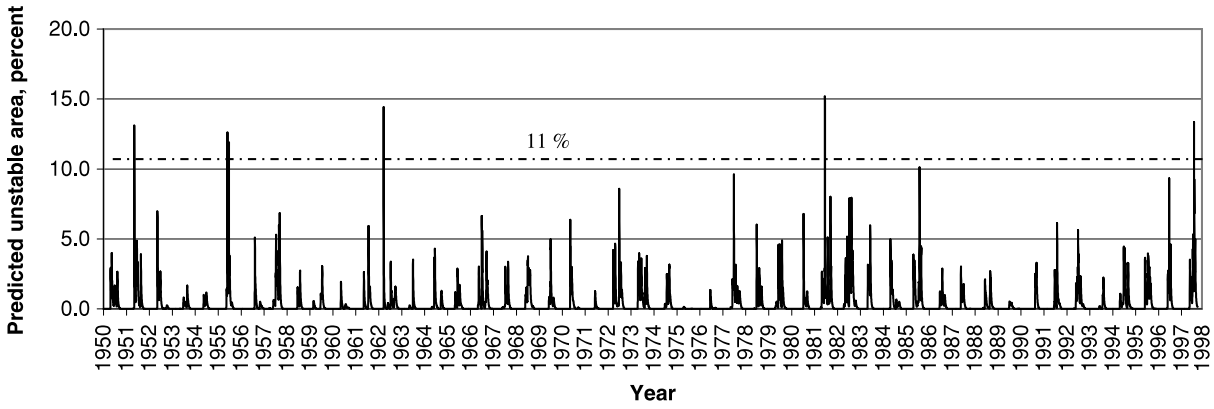

Figure 11. Model instability prediction calculated over the 1950–98 period using the best-fit simulation. Using an empirical threshold of 11% for the catchment, the model predicts the three landslide-producing events (1955, 1982 and 1998), but also predicts two false positives (1951–52 and 1962–63)

582 mm for 13 February 1986, computed from 1 July). Nonetheless, the 1986 event had no significant effect on hillslope stability at the study site.

The parameter combination $K_1 = 50$ m day^{-1}, $K_2 = 2$ m day^{-1}, $f_1 = 2$ m^{-1}, and $f_2 = 2$ m^{-1} (simulation no. 40 in Figure 9) was the one that detected better the debris flow event of January 1982 (e.g. the only peak exceeding the 11% threshold) and was chosen as 'best-fit'; other simulations incorrectly predicted other landslide events. When applied to the entire time series, the best-fit model predicted correctly all three landslide events (Figure 11), based on the criterion of the threshold of 11% of the catchment area as unstable. The corresponding maps of the factor of safety are shown in Figure 12. Two storms, however, are predicted to produce landslides, but they did not. Furthermore, the spatial extent of landsliding was overpredicted for all three events. Below, we discuss the possible explanations of this incorrect prediction.

DISCUSSION

Our calibrated mechanistic model performed better than the simple threshold-based *I–D* curves. Using the conditions that predicted instability area must exceed 11% of the catchment to be a landslide-producing storm, only two storms (in the 1951–52 and 1962–63 hydrological years) are incorrectly predicted as unstable. We explore four basic causes for incorrectly predicting landslides-producing storms once a calibration has occurred. These are potential sources of error for any model.

Daily rainfall data is too coarse a time scale

It has been observed (e.g. Johnson and Sitar, 1990; Reid *et al*., 1997; Montgomery *et al*., 1997; Torres *et al*. 1998) that a significant fraction of hillslope failures is often related to short (≤1 h) and intense rainfall rather than daily-averaged precipitation. To explore whether our model would produce significantly different outcomes based on hourly rainfall, we generated some simulations using six synthetic hourly hietographs (Figure 13). Every synthetic hietograph has the same daily totals as the 4 and 5 January 1982 observed daily rainfall, but different hourly distributions, ranging from constant moderate rainfall to high rainfall peaks. The hietograph no. 6 was normalized from the representative rainfall curve for northern California (Ellen and Wieczorek, 1988). We compared the pore pressure responses calculated from the different storm inputs, for convergent areas (hollows) and divergent areas (noses). The timing of the calculated peak pore pressure differed from 4 to 9 h. The magnitude of the calculated peak pore pressure response differed little among the six runs. Hence, although the difference in hourly rainfall sequence changed slightly the timing of the peak pore pressure response, it had little effect on its magnitude, and the use of hourly rainfall record would not improve substantially the performance of this model. It is possible that a model that can route pressure waves generated for short rainfall bursts (e.g. Marui *et al*., 1993; Torres *et al*., 1998; Iverson, 2000) would provide a more reliable prediction. If such a model were to be developed

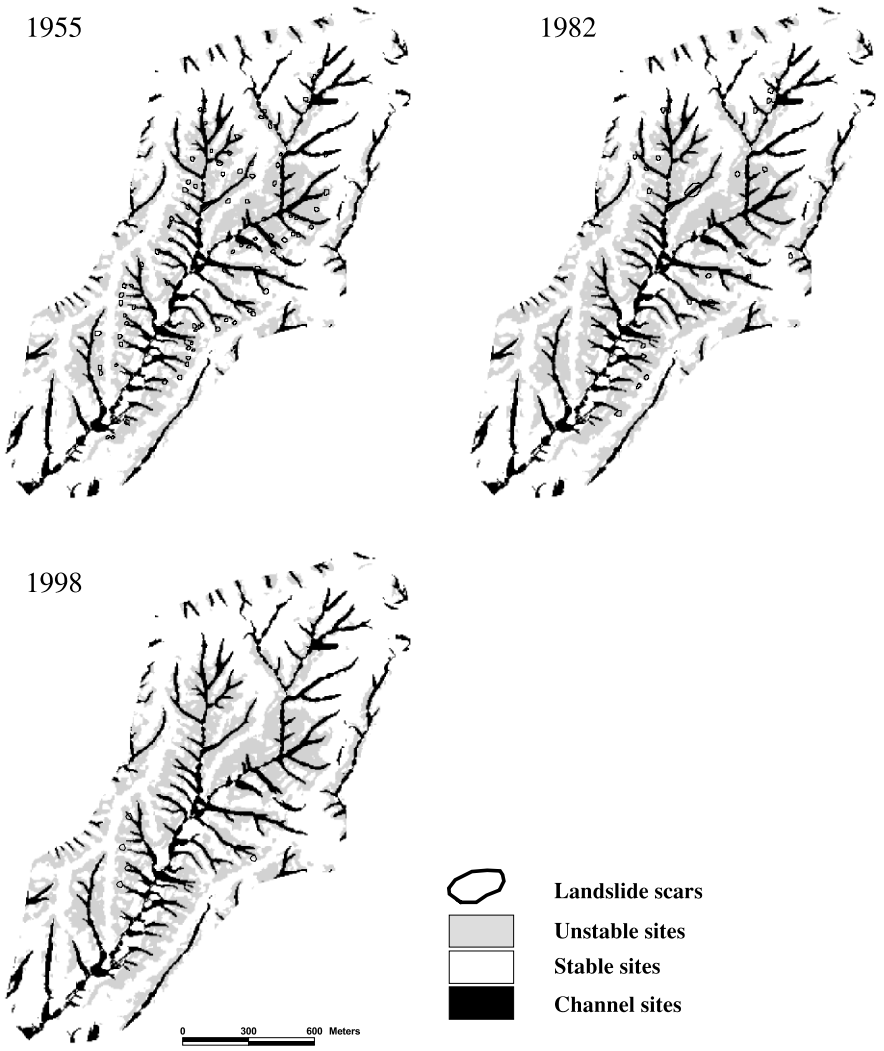

Figure 12. Map of the factor of safety calculated by the best-fit model at the peak of the three debris-flow-producing storms

that could run continuously at high-resolution time intervals (minutes), it would currently outstrip computational capabilities when applied as a forecasting tool to large areas. Furthermore, we currently lack either the observational data or theoretical prediction at suffiently high spatial and temporal resolution to use such a model.

Rainfall data are not sufficiently local

The closest rainfall record is located about 6 km south of the catchment. Although the hietograph map by Monteverdi (1984) suggests that, on average, the rainfall record from the Half Moon Bay station may be a reasonable proxy, it is likely to deviate more if the short-term input becomes the primary triggering factor. This issue is even more important for the use of the hourly time series from the San Francisco International Airport.

Previous landslides create legacy effects

The number of observed landslides in our study site decreased in time regardless of the rainfall pattern, and the greatest number of debris flows events did not correspond to the most intense storms (Figure 4). This pattern suggests that there may be a legacy effect, i.e. the interval between potential landslide-producing storms is short relative to the recovery time of the most susceptible areas to instability. A shallow landslide typically removes

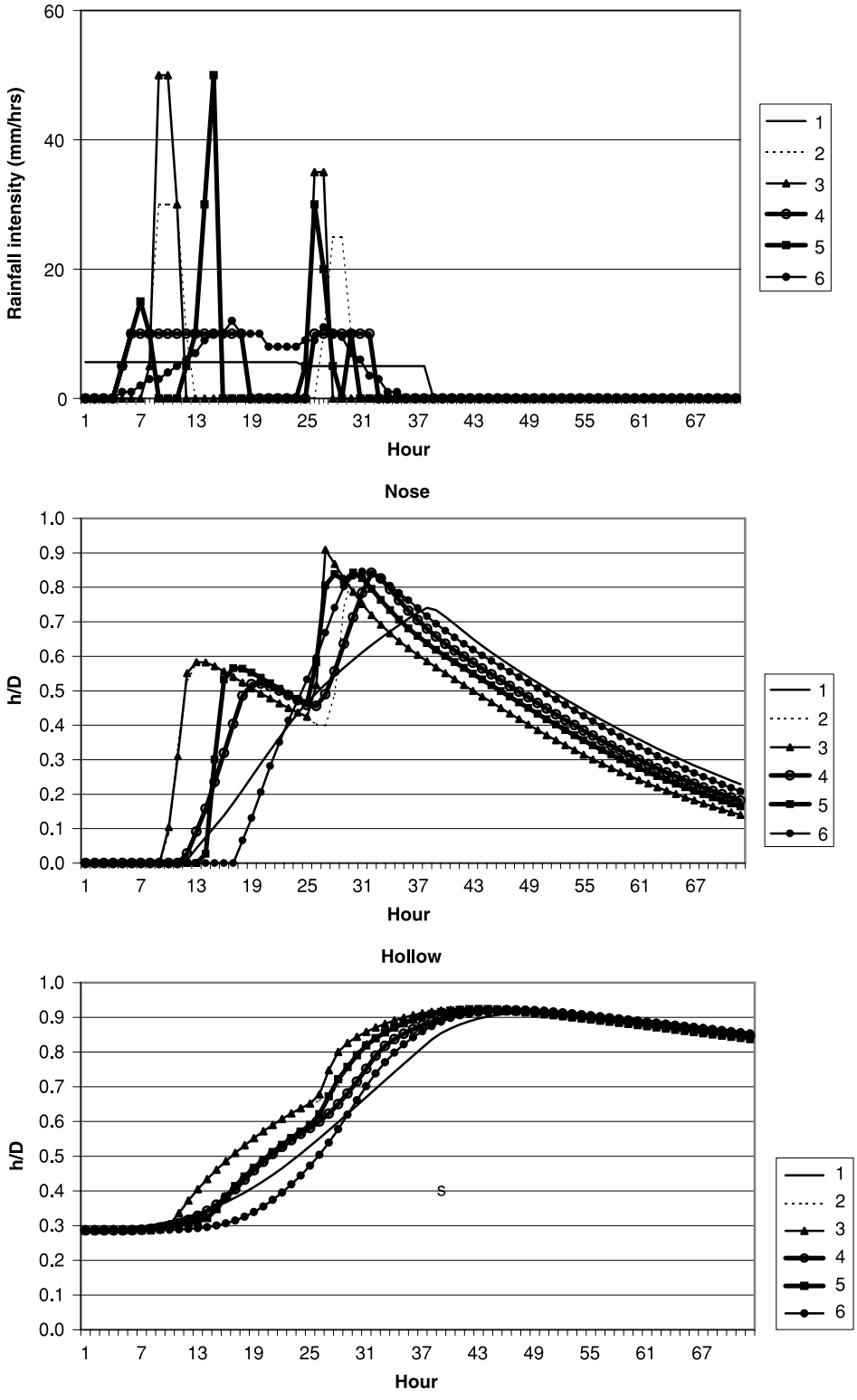

Figure 13. Synthetic hourly rainfall input to the model for the 3–5 January 1982 event, and the calculated piezometric response at a topographically convergent site (hollow) and at a divergent site (nose). At each site the magnitude of the peak response is nearly identical, and the delay between maximum peaks ranges between 4 and 9 hours

the entire soil mantle and creates a well-drained scar into the hillslope. Stratigraphic studies (Dietrich and Dunne, 1978; Dietrich *et al*., 1982), colluvium dating (Reneau and Dietrich, 1986, 1987) and modelling (Reneau, 1988; Hsu, 1994; Dietrich *et al*., 1995) indicate that recovery time of soil thickness in typical failure sites in California is thousands of years. Crozier and Preston (1999) argued for a legacy effect in describing the landslide response in New Zealand to recent deforestation. They argued that landslides reduce available soil thickness and leave a strong mass less susceptible to failure. Hence, initial massive instability resulting from deforestation is followed by a drop in landslide frequency for a given storm. They proposed, however, that, in time, sufficient site recovery would lead to renewed site susceptibility to failure. They specifically argued, therefore, that rainfall thresholds would change through time due to the legacy effect. When this is the case, it is worth emphasizing that this problem will be overlooked by any model calibrated over too short a time period; for example, in our data set, the rainfall threshold approach works reasonably well for the 1970–86 period, but the inclusion of the precipitation and landslides occurring in the 1950–70 period shows a poor performance that we argue could be explained by the legacy effect.

Accounting for the legacy effect, while clearly important, did not, however, improve our modelling results. The total area that actually fails relative to that predicted as unstable over the catchment is very small. Consequently, eliminating the cells that failed in 1955 from the simulation of the following years did not eliminate the problem of predicted 'false positives'. This suggests that a dominant control on model performance may be the inaccurate topography and soil properties.

Topography and soil properties are inaccurate

In steep terrain, 5–10 m grid DEMs derived from the US Geological Survey 7·5′ maps will generally underestimate the local slope steepness that controls shallow landsliding. Although the topographic data captured part of the valley network and the hollows from which landslides most commonly occurred (Figure 14), many fine-scale valleys are missing, especially the first-order channels that are common sources of debris flows. Furthermore, the data generally do not capture local steps and gentle areas, nor the strength of the topographic

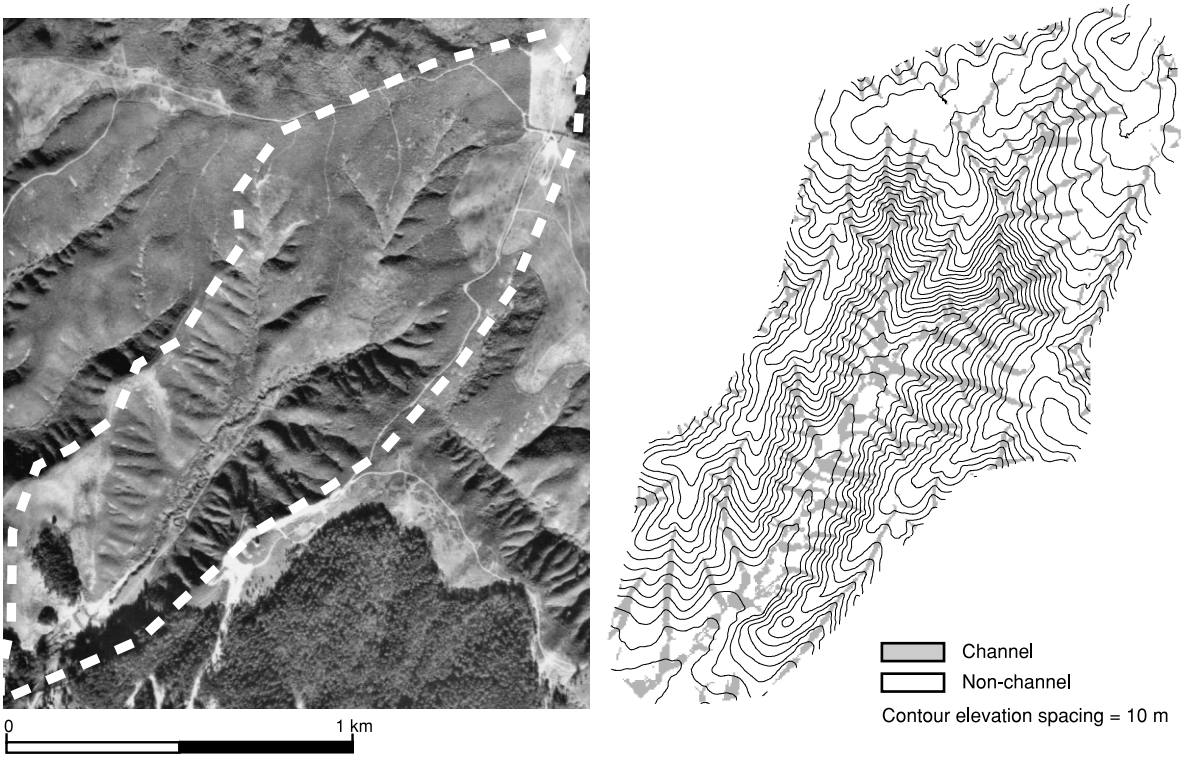

Figure 14. Comparison of the actual valley network and at derived from the DEM using a drainage area threshold (1000 m^2)

convergence and divergence, which influence pore pressure build up due to shallow subsurface flow (e.g. Dietrich *et al*., 2001). Even with high-resolution data derived from airborne laser altimetry, there is a tendency for topographically driven landslide models to overpredict actual landslide occurrence (Dietrich *et al*., 2001).

CONCLUSIONS

The increasing ability to forecast precipitation accurately and the increasing resolution of digital elevation data invites modelling efforts that can be used to predict the timing and location of shallow landslide hazards. In order to be of practical value, such models must not miss significant landslide-producing storms, but neither should they frequently overpredict landslide occurrence. The current practice of using *I–D* thresholds to identify landslide-producing storms can be made more useful if they are coupled with mechanistic models that are spatially explicit in identifying areas of high risk.

Our result suggests that landslide prediction based on dynamic hydrological models may perform significantly better than simple *I–D* approaches, but they require considerable calibration. It is likely that calibration parameters will be relatively site specific. The paucity of time series data on landslide occurrence will make it difficult to calibrate such models over large areas. Furthermore, poor topography, precipitation, soil properties data and legacy effects of previous landslides will inhibit model performance. The *I–D* model coupled with a tool for delineating the spatial extent of landslide risk can be easily assembled but will have a large uncertainty. Further improvements in high-resolution precipitation and topographic data, coupled with the inventory of landslides time series data, should allow more mechanistic models to be calibrated and provide significant improvement in hazard forecasting (e.g. Dietrich *et al*., 2001).

We developed a dynamic distributed hydrological–slope stability model based on daily rainfall and topographic data. The application of the model showed that the calibration yields both spatial and temporal overprediction of instability.

By selecting an appropriate threshold for predicted unstable area, however, the model performed better than previous rainfall-based models in discriminating landslide-producing storms. The overprediction of instability could not be removed by the use of hourly rainfall data and neither did such data affect significantly the worst-case response calculated by our model, probably because the advective mechanism used for water recharge is too slow to reproduce correctly the pore-pressure response to high-intensity rainfall; this is true for most past and present distributed hillslope hydrology and slope stability models.

The incorporation of the 'legacy effect' was not able to remove the overprediction either. The removal of soil mantle from the scars was incorporated in the model, but the results for the 1982 and 1986 events were not significantly different.

The above approach will give better results by improving the quality and detail of topographic data, rainfall data and ground characterization, especially in terms of soil thickness. The wider availability of hourly rainfall data will allow calibration of more sophisticated physical models for the unsaturated pore pressure response (e.g. Iverson 2000).

In terms of practical applications, however, we believe that overprediction of instability will be inevitable for a warning system operating in ungauged catchments, and further efforts should be made to increase the awareness of this in the communities subject to debris flow hazard.

ACKNOWLEDGEMENTS

This research is funded by NASA under the 'Cassandra' project, NRA-98-OES-13, fund no. 23944. The work used the data on topography and landslides digitized by Mei-Ling Hsu for her PhD dissertation. Dino Bellugi ran the model for soil depth estimation and helped at various stages with discussions and suggestions.

REFERENCES

Beven KJ. 2001. *Rainfall-Runoff Modelling. The Primer*., John Wiley and Sons Ltd.

Beven KJ, Kirkby MJ. 1979. A physically based variable contributing area model of catchment hydrology. *Hydrological Sciences Bulletin* **24**: 43–69.

Brooks RH, Corey WT. 1964. *Hydraulic properties of porous media*. Hydrology Paper **3**, Colorado State University, Fort Collins.
Caine N. 1980. The rainfall intensity–duration control of shallow landslides and debris flows. *Geografisker Annaler, Series A* **62**: 23–27.
Campbell RH. 1975. Soil slips, debris flows and rainstorms in the Santa Monica Mountains and vicinity, southern California. *United States Geological Society Professional Paper* **851**.
Cancelli A, Nova R. 1985. Landslides in soil debris covertriggered by rainstorms in Valtellina (Central Alps – Italy). In *Proceedings 4th International Conference on Landslides, Tokyo*. The Japan Geological Society: 267–272.
Cannon SH, Ellen S. 1985. Rainfall conditions for abundant debris avalanches. San Francisco Bay Region, California. *California Geology* **38**: 267–272.
Canuti P, Focardi P, Garzonio CA. 1985. Correlation between rainfall and landslides. *Bulletin of the International Association of Engineering Geologists* **31**: 49–54.
Carrara A. 1983. Multivariate models for landslide hazard evaluation. *Mathematical Geology* **15**: 403–426.
Carrara A. 1993. Uncertainty in evaluating landslide hazard and risk. In *Prediction and Perception of Natural Hazards*, Nemec J, Nigg JM, Siccardi F (eds). Kluwer Academic Publishers: Dordrecht; 101–109.
Carrara A, Cardinali M, Detti R, Guzzetti F, Pasqui V, Reichenbach P. 1991. GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Processes and Landforms* **16**: 427–445.
Carrara A, Cardinali M, Guzzetti F, Reichenbach P. 1995. GIS technology in mapping landslide hazard. In *Geographical Information System in Assessing Natural Hazard*, Carrara A, Guzzetti F (eds). Kluwer: New York; 135–176.
Chow VT, Maidment DR, Mays LW. 1988. *Applied Hydrology*. McGraw Hill.
Chung CF, Fabbri AG, Van Westen CJ. 1995. Multivariate regression analysis for landslide hazard zonation. In *Geographical Information System in Assessing Natural Hazard*, Carrara A, Guzzetti F (eds). Kluwer: New York; 107–134.
Church M, Miles MJ. 1987. Meteorological antecedent to debris flows in southwestern British Columbia; some case studies. *Geological Society of America, Reviews in Engineering Geology* **7**: 63–79.
Costa Cabral MC, Burges SJ. 1994. Digital elevation model networks (DEMON): a model for flow over hillslopes for computation of contributing and dispersal areas. *Water Resources Research* **30**(6): 1681–1692.
Crosta G. 1998. Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology* **35**(2–3): 131–145.
Crosta G, Marchetti M. 1993. Frane superficiali nel medio tratto della Valle Olona. *Geologia Applicata e Idrogeologia* **32**: 60–77.
Crozier MJ. 1999. Prediction of rainfall-triggered landslides: a test of the antecedent water status model. *Earth Surface Processes and Landforms* **24**: 825–833.
Crozier MJ, Preston NJ. 1999. Modelling changes in terrain resistance as a component of landform evolution in unstable hill country. In *Process Modelling and Landform Evolution*, Hergarten S, Neugebauer HJ (eds). Lecture Notes in Earth Science, Vol. 78. Springer-Verlag: Heidelberg; 267–284.
Delecluse P. 1999. El Nino et sa prevision (El Nino and its predictability). In *L'Effet de serre, certitudes et incertitudes. Montrouge: Gauthier-Villars.* Petit M (ed.). *Comptes Rendus de l'Academie des Sciences, Serie II. Sciences de la Terre et des Planetes* **328**(4): 281–288.
Dietrich WE, Dunne T. 1978. Sediment budget for a small catchment in mountainous terrain. *Zeitschmft für Geomorphologie, Supplementband* **29**: 191–206.
Dietrich WE, Dunne T, Humphrey NF, Reid LM. 1982. Construction of sediment budgets for drainage analysis. In *Sediment budgets and routing in forested basins*, Swanson FJ, Janda RJ, Dunne T, Swanston DN (eds). Forest Service USDA General Technical Report, **PNW-141**; 5–23.
Dietrich WE, Wilson CJ, Montgomery DR, MacKean J, Bauer R. 1992. Erosion thresholds and land surface morphology. *Geology* **20**: 675–679.
Dietrich WE, Wilson CJ, Montgomery DR, MacKean J. 1993. Analysis of erosion thresholds, channel networks and landscape evolution using a digital terrain model. *Journal of Geology* **101**: 259–278.
Dietrich WE, Reiss R, Hsu M, Montgomery DR. 1995. A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. *Hydrological Processes* **9**: 383–400.
Dietrich WE, Bellugi D, Real De Asua R. 2001. Validation of the shallow landslide model, SHALSTAB, for forest management. In *Land Use and Watersheds: Human Influence on Hydrology and Geomorphology in Urban and Forest Areas*, Wigmosta MS, Burges SJ (eds). Water Science and Application **2**. American Geophysical Union: 195–227.
Division of Water Resources, Department of Public Works, State of California. 1956. *Floods of December 1955 in California*.
Duan J. 1996. *A coupled hydrologic–geomorphic model for evaluating effects of vegetation change in watersheds*. PhD dissertation, Oregon State University.
Ellen SD, Wieczorek GK. 1988. Landslides, floods and marine effects of the storm of January 3–5, 1982 in the San Francisco Bay region, California. *United States Geological Survey Professional Paper* **1434**.
Freeman TJ. 1991. Calculating catchment area with divergent flow based on a regular grid. *Computers and Geosciences* **17**: 412–422.
Glade T. 1997. Establishing the frequency and magnitude of landslide-triggering rainstorm events in New Zealand. *Environmental Geology*, **35**(2–3): 160–174.
Glade T, Crozier M, Smith P. 2000. Applying probability determination to refine landslide triggering rainfall thresholds using an empirical 'Antecedent daily rainfall model'. *Pure and Applied Geophysics* **157**: 1059–1079.
Grayson RB, Moore ID, Smith P. 1992. Physically-based hydrologic modelling: I. a terrain-based model for investigative purposes. *Water Resources Research* **28**(10): 2639–2658.
Haith DA, Shoemaker LL. 1987. Generalized watershed loading functions for stream flow nutrients. *Water Research Bulletin* **23**: 471–478.
Hamon WR. 1961. Estimating potential evapotranspiration. *Proceedings of the American Society of Civil Engineers, ASCE Proceedings Journal of the Hydraulics Division* **87**(HY3): 107–120.
Heimsath AM, Dietrich WE, Nishiizumi K, Finkel RC. 1997. The soil production function and landscape equilibrium. *Nature* **388**: 358–361.
Heimsath AM, Dietrich WE, Nishiizumi K, Finkel RC. 1999. Cosmogenic nuclides, topography and the spatial variation of soil depth. *Geomorphology* **27**: 151–172.
Hsu M. 1994. *A grid-based model for predicting soil depth and shallow landslides*. PhD dissertation, U.C. Berkeley.

Iverson RM. 1990. Groundwater flow fields in infinite slopes. *Geotechnique* **40**: 139–143.
Iverson RM. 2000. Landslide triggering by rain infiltration. *Water Resources Research* **36**(7): 1897–1910.
Johnson KA, Sitar N. 1990. Hydrologic conditions leading to debris flow initiation. *Canadian Geotechnical Journal* **27**: 789–801.
Keefer DK, Wilson RC, Mark RK, Brabb EE, Brown W, Ellen SD, Harp EL, Wieczorek GF, Alger CS, Zatkin RS. 1987. Real-time landslide warning during heavy rainfall. *Science* **238**: 921–925.
Kim SK, Hong WP, Kim YM. 1992. Prediction of rainfall-triggered landslide in Korea. In *Proceedings of 6th International Symposium on Landslides, Christchurch*, Bell D (ed.). Balkema: Rotterdam; 989–994.
Lamb R, Beven KJ, Myrabo S. 1997. Discharge and water table predictions using a generalized TOPMODEL formulation. *Hydrological Processes* **11**: 1145–1167.
Lamb R, Beven KJ, Myrabo S. 1998. Use of spatially distributed water observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources* **22**(4): 305–317.
Larsen MC, Simon A. 1993. A rainfall intensity–duration threshold for landslides in a humid–tropical environment. *Geografiska Annaler Series A* **75**: 13–23.
Marui A, Yasuhara M, Kuroda K, Takayama S. 1993. Subsurface water movement and transmission of rainwater pressure through a clay layer. In *Hydrology of Warm Humid Regions*, Gladwell JS (ed.). IAHS Publication 216. IAHS Press: Wallingford; 463–470.
Miller NL, Kim J. 1997a. The regional climate system model. In *Mission Earth: Modeling and Simulation for a Sustainable Global System.* Clymer M, Mechoso C (eds). SIAM: chapter 19.
Miller NL, Kim J. 1997b. Numerical prediction of precipitation and riverflow over the Russian River watershed during the January 1995 California flooding. *Bulletin of the American Meteorological Association* **77**: 101–105.
Monteith JL. 1965. Evaporation and environment. *Symposia of the Society for Experimental Biology* **19**: 205–234.
Monteverdi JL. 1984. *Debris flows, landslides and floods in the San Francisco Bay region, January 1982*. Overview and summary of a conference held at Stanford University August 23–26, 1982. National Academy Press, Washington.
Montgomery DR. 1991. *Channel initiation and landscape evolution*. PhD dissertation, U.C. Berkeley.
Montgomery DR, Dietrich WE. 1994. A physically based model for the topographic control on shallow landsliding. *Water Resources Research* **30**(4): 1153–1171.
Montgomery DR, Dietrich WE, Torres R, Anderson SP, Loague K. 1997. Subsurface flow paths in a steep unchannelled catchment. *Water Resources Research* **33**(1): 91–109.
Montgomery DR, Schmidt KM, Green Berg HM, Dietrich WE. 2000. Forest clearing and regional landsliding. *Geology* **28**(4): 311–314.
Moore ID, O'Loughlin EM, Burch GJ. 1988. A contour based topographic modeland its hydrological and ecological applications. *Earth Surface Processes and Landforms* **13**: 305–320.
Nilsen T, Turner BL. 1975. Influence of rainfall and ancient landslide deposits on recent landslides (1950–71) in urban areas of Contra Costa County, California. *United States Geological Society Bulletin* **1388**.
Okimura T. 1989. Prediction of slope failure using the estimated depth of the potential failure layer. *Journal of Natural Disaster Science* **11**(1): 67–79.
Okimura T, Kawatani T. 1987. Mapping of the potential surface-failure sites on granite slopes. In *International Geomorphology 1986, Part I*, Gardiner V (ed.). Wiley: Chichester; 121–138.
Pack RT, Tarboton DG, Goodwin CN. 1998. *Terrain stability mapping with SINMAP, technical description and users guide for version 1.00*, Report Number 4114-0, Terratech Consulting Ltd, Salmon Arm, B.C., Canada.
Parlange JY, Steenhuis TS, Haverkamp R, Barry DA, Culligan PJ, Hogarth WL, Parlange MB, Ross P, Stagnitti F. 1999. Soil properties and water movement. In *Vadose Zone Hydrology. Cutting Across Disciplines*, Parlange MB, Hopmans JW (eds). Oxford University Press: New York, Oxford.
Penman HL. 1948. Natural evaporation from open water, bare soil, and grass. *Proceedings of the Royal Society London*, *Series A* **193**: 120–195.
Piechota TC, Dracup JA, Fovell RG. 1997. Western US streamflow and atmospheric circulation patterns during El Nino–southern oscillation. *Journal of Hydrology* **201**(1–4): 249–271.
Pradel D, Raad G. 1993. Effect of permeability on surficial stability of homogeneous slopes. *Journal of Geotechnical Engineering ASCE* **119**: 315–332.
Quinn PF, Beven KJ, Chevalier P. 1991. The prediction of hillslope flowpaths for distributed modelling using digital terrain models. *Hydrological Processes* **5**: 59–80.
Rahardjo H, Lim TT, Chang MF, Fredlund DG. 1995. Shear strength characteristics of a residual soil. *Canadian Geotechnical Journal* **32**: 60–77.
Reid ME, LaHunsen RG, Iverson RM. 1997. Debris flow initiation experiments using diverse hydrologic triggers. In *Debris Flow Hazards Mitigation: Mechanics, Prediction and Assessment*, Chen CL (ed.). ASCE Proceedings. ASCE: 1–11.
Reneau SL. 1988. *Depositional and erosional history of hollow: application to landslide location and frequency, long-term erosion rates and the effect of climatic change*. PhD dissertation, U.C. Berkeley.
Reneau SL, Dietrich WE. 1986. Size and location of colluvial landslides in a steep forested landscape. Proceedings of International Symposium on Erosion and Sedimentation in the Pacific Rim, 3–7 August 1987, Corvallis, Oregon. *International Association of Hydrological Sciences Bulletin* **165**: 39–48.
Reneau SL, Dietrich WE. 1987. The importance of hollows in debris flow studies; examples from Marin County, California. *Geological Society of America Reviews in Engineering Geology* **7**: 165–180.
Schmidt KM, Roering JJ, Stock JD, Dietrich WE, Montgomer DR, Schaub TL. 2001. The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range. *Canadian Geotechnical Journal* **38**(5): 995–1024.
Seo DJ, Breidenbach JP, Johnson ER. 1999. Real-time estimation of mean field bias in radar rainfall data. *Journal of Hydrology* **223**(3–4): 131–147.
Skempton AW, DeLory FA. 1957. Stability of natural slopes in London clay. ASCE Journal 2: 378–381.
Sidle RC. 1992. A theoretical model of the effects of timber harvesting on slope stability. *Water Resources Research* **28**(7): 1897–1910.
Terlien MTJ. 1997. The determination of statistical and deterministic hydrological landslide-triggering thresholds. *Environmental Geology* **35**(2–3): 124–130.

Terlien MTJ, Asch TWJ, Van Westen CJ. 1995. Deterministic modelling in GIS-based landslide hazard assessment. In *Geographical Information System in Assessing Natural Hazard*, Carrara A, Guzzetti F (eds). Kluwer: New York, 57–77.
Torres R, Dietrich WE, Montgomery DR, Anderson SP, Loague K. 1998. Unsaturated zone processes and the hydrologic response of a steep, unchannelled catchment. *Water Resources Research* **34**(8): 1865–1879.
Tucker GE, Slingerland RL. 1997. Drainage basin responses to climate change. *Water Resources Research* **33**: 2031–2047.
Van der Werf B. 1992. *Granada, a Synonym of Paradise*. Gum Tree Lane Books: El Granada, CA.
Van Westen CJ. 1993. *Application of Geographic Information Systems to Landslide Hazard Zonation*. ITC Publication **15**. ITC: Enschede.
Vicente GA, Scofield RA. 1998. Satellite rainfall estimates in real time for applications to flash flood watches and warnings, heavy precipitation forecasting and asimilation of numerical weather prediction models. In *16th Conference on Weather Analysis and Forecasting*, American Meteorological Society, Minneapolis; 790–793.
Wagner RJ, Nelson RE. 1961. *Soil Survey of the San Mateo Area, California*. USDA, Soil Conservation Service, series 1954, no. 13. USDA, Soil Conservation Service.
Wentworth CM. 1986. Maps of debris flow features evident after the storms of December 1955 and January 1982, Montara Mountains area, California. *USGS Open-File Map* **86–363**.
Wieczorek GF. 1987. Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains, California. *Geological Society of America Reviews in Engineering Geology* **7**: 93–104.
Wigmosta MS, Lettenmaier DP. 1999. A comparison of simplified methods for routing topographically driven subsurface flow. *Water Resources Research* **5**(1): 255–264.
Wilson CJ. 1988. *Runoff and pore pressure development in hollows*. PhD dissertation, U.C. Berkeley.
Wilson RC. 1997. Broad-scale climatic influences on rainfall thresholds for debris flows: adapting thresholds for northern California to southern California. *Geological Society of America Reviews in Engineering Geology* **11**: 71–80.
Wilson RC. 2000. Climatic variations in rainfall thresholds for debris-flow activity. In *Proceedings of First Plinius Conference on Mediterranean Storms*, Mareta, Italy, 14–16 October 1999, Claps P, Wieczorek GW (eds). European Geophysical Union: 415–442.
Wilson RC, Jayko AS. 1997. *Map of Rainfall Thresholds for Debris Flows in the San Francisco Bay Region, California*. USGS Open File Report 97–745 F, Maps 1:275, 000. http://wrgis.wr.usgs.gov/open–file/of97–745.
Wolock DM. 1993. *Simulating the variable-source-area concept of streamflow generation with the watershed model TOPMODEL*. USGS Water-Resources Investigations WRI textbf93-4124.
Wu W. 1993. Distributed slope stability analysis in steep forested basins. PhD dissertation, Utah State University.
Wu W, Sidle R. 1995. A distributed slope stability model for steep forested basins. *Water Resources Research* **31**: 2097–2110.

